University of Alberta

Decision Points: Determining Factors in Critical Appraisal of Health-Related Web Resources

by

Jason Scott Daniels

©

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Psychology

Edmonton, Alberta
Fall 2007

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-32949-8
Our file Notre référence
ISBN: 978-0-494-32949-8

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Critical appraisal is a skill that is becoming more important as the World Wide Web (Web) continues to influence the ways in which information is accessed. This is especially important in health related domains where public access to medical journals and health-related information, both accurate and inaccurate, has also increased tremendously. Despite this fact, it seems that people are either unwilling or unable to effectively evaluate health related information they find on the Web. I propose a model to examine the role of factors influencing the likelihood of critical appraisal and to establish a framework for future research in this area. I determined that both ability factors and motivational factors are important in determining whether critical appraisal will occur. Specifically, ability factors such as domain knowledge, knowledge and skill with the Web, and knowledge of critical appraisal influence whether critical appraisal can occur and motivational factors determine whether critical appraisal will occur.

Table of Contents

DECISION POINTS: DETERMINING FACTORS IN CRITICAL APPRAISAL OF HEALTH-RELATED WEB RESOURCES 1

The Development of Literacy 3

History of the Web 6

Critical Appraisal 9

The Consequences of Using Information from the Web 11

Health Information and the Web 12

The Draw of the Web 14

Assessing the Quality of Online Health Information 17

Factors Influencing Critical Appraisal 18

Domain Knowledge 19

Knowledge and Skill with the Web 23

Knowledge of Critical Appraisal 23

Motivation 24

A Modified Approach 27

Domain Knowledge 29

Knowledge and Skill with the Web 29

Knowledge of Critical Appraisal 30

Motivation 31

Critical Appraisal 31

Structural Equation Modeling 33

STUDY 1: ESTABLISHMENT AND RE-SPECIFICATION OF THE

BASELINE MODEL 36
  Methods 36
    Participants 36
    Measures and Scoring 36
      Domain knowledge 36
      Knowledge and skill with the Web 37
      Knowledge of critical appraisal 37
      Motivation 38
      Critical appraisal 38
    Procedure 40
  Results 41
    Specification of Baseline Model 41
      Descriptive statistics 41
        Domain knowledge 41
        Knowledge and skill with the Web 44
        Knowledge of critical appraisal 48
        Motivation 56
        Critical Appraisal 61
      Structural analysis 66
    Re-specification of the Baseline Model 76
    Analysis of the Impact of Covariance Measures on the Structural Model 84

Analysis of the Impact of Motivation as a Moderating Variable 90
Discussion 96
STUDY 2: CROSS-VALIDATION OF RE-SPECIFIED MODEL 107
Methods 108
Participants 108
Measures and Scoring 108
Procedure 108
Collection of validation sample 108
Cross-validation 108
Results 109
Descriptive Statistics 109
Domain knowledge 109
Knowledge and skill with the Web 114
Knowledge of critical appraisal 118
Motivation 126
Critical appraisal 131
Omnibus Test 136
Constrained Test 136
Cross-validations Measure 136
Discussion 141
GENERAL DISCUSSION 144
REFERENCES 159

APPENDIX A 172
APPENDIX B 179
APPENDIX C 180
APPENDIX D 182
APPENDIX E 187
APPENDIX F 189
APPENDIX G 193
APPENDIX H 200

List of Tables

Table 1
Standardized Regression Weights of Parameter Estimates for Latent Variables (Full Baseline Model) 72

Table 2
Summary of Fit Indices for Full Baseline Model 73

Table 3
Covariance Measures between Latent Variables (Full Baseline Model) 75

Table 4
Standardized Regression Weights of Parameter Estimates for Latent Variables (Re-specified Baseline Model) 79

Table 5
Summary of Fit Indices for Re-specified Baseline Model 81

Table 6
Covariance Measures between Latent Variables (Re-specified Baseline Model) 83

Table 7
Standardized Regression Weights of Parameter Estimates for Latent Variables (No-Covariance Baseline Model) 87

Table 8
Summary of Fit Indices for No-Covariance Baseline Model 88

Table 9
Standardized Regression Weights of Parameter Estimates for Latent Variables (Motivation Covariance Measures Removed) 93

Table 10
Summary of Fit Indices for Baseline Model (Motivation Covariance Measures Removed) 94

Table 11
Independent-Sample t-tests Between Baseline Measures and Re-specified Baseline Model 110

Table 12
Summary of Fit Indices for the Omnibus Cross Validation 137

Table 13
Summary of Fit Indices for the Constrained Cross Validation 139

List of Figures

Figure 1.
A proposed model of the factors related to critical appraisal of health related Web resources. 20

Figure 2.
A model of the possible interactions between factors related to critical appraisal of health related Web resources and critical appraisal. 28

Figure 3.
A histogram of the frequency of test of alcohol abuse domain knowledge (baseline sample). 42

Figure 4.
A bar graph of the frequency of self-reported domain knowledge (baseline sample). 43

Figure 5.
A histogram of the frequency of participants' ratings of how much they like to use Computers and the Web (baseline sample). 45

Figure 6.
A histogram of the frequency of participants' ratings of how frequently they use Computers and the Web (baseline sample). 46

Figure 7.
A histogram of the frequency of participants' reported skill with Computers and the Web (baseline sample). 47

Figure 8.
A histogram of participants' strategies for finding the Author of a Web resource (baseline sample). 49

Figure 9.
A histogram of participants' strategies for assessing the credibility or believability of the author of a Web resource (baseline sample). 50

Figure 10.
A histogram of participants' strategies for determining the purpose of a Web resource (baseline sample). 51

Figure 11.
A histogram of participants' strategies for assessing the accuracy of a Web resource (baseline sample). 52

Figure 12.
A histogram of participants' strategies for assessing the objectivity of a Web resource (baseline sample). 53

Figure 13.
A histogram of participants' strategies for assessing the comprehensiveness of a Web resource (baseline sample). 54

Figure 14.
A histogram of participants' strategies for assessing the currency of a Web resource (baseline sample). 55

Figure 15.
A histogram of participants' ratings of personal relevance (baseline sample). 57

Figure 16.
A histogram of participants' ratings of personal responsibility (baseline sample). 58

Figure 17.
A bar graph of participants' ratings of their motivation on the day of testing (baseline sample). 59

Figure 18.
A bar graph of participants' ratings of their general motivation to ensure that information regarding alcohol abuse is accurate (baseline sample). 60

Figure 19.
A bar graph of participants' ratings of the test Web resource (baseline sample). 62

Figure 20.
A histogram of participants' critical appraisal scores regarding the author of a Web resource (baseline sample). 63

Figure 21.
A histogram of participants' critical appraisal scores regarding the content of a Web resource (baseline sample). 64

Figure 22.
A histogram of participants' critical appraisal scores regarding the form of a Web resource (baseline sample). 65

Figure 23.
A proposed structural model including latent variables. 70

Figure 24.
A proposed structural model including latent variables and standardized regression weights. 71

Figure 25.
Proposed re-specified structural model. 78

Figure 26.
Re-specified structural model with standardized regression weights. 80

Figure 27.
Baseline model with all covariance measures removed. 86

Figure 28.
Baseline model with motivation covariance measures removed. 92

Figure 29.
A histogram of the frequency of test of alcohol abuse domain knowledge (validation sample). 112

Figure 30.
A bar graph of the frequency of self-reported domain knowledge (validation sample). 113

Figure 31.
A histogram of the frequency of participants' ratings of how much they like to use Computers and the Web (validation sample). 115

Figure 32.
A histogram of the frequency of participants' ratings of how frequently they use Computers and the Web (validation sample). 116

Figure 33.
A histogram of the frequency of participants' reported skill with Computers and the Web (validation sample). 117

Figure 34.
A histogram of participants' strategies for finding the Author of a Web resource (validation sample). 119

Figure 35.
A histogram of participants' strategies for assessing the credibility or believability of the author of a Web resource (validation sample). 120

Figure 36.
A histogram of participants' strategies for determining the purpose of a Web resource (validation sample). 121

Figure 37.
A histogram of participants' strategies for assessing the accuracy of a Web resource (validation sample). 122

Figure 38.
A histogram of participants' strategies for assessing the objectivity of a Web resource (validation sample). 123

Figure 39.
A histogram of participants' strategies for assessing the comprehensiveness of a Web resource (validation sample). 124

Figure 40.
A histogram of participants' strategies for assessing the currency of a Web resource (validation sample). 125

Figure 41.
A histogram of participants' ratings of personal relevance (validation sample). 127

Figure 42.
A histogram of participants' ratings of personal responsibility (validation sample). 128

Figure 43.
A bar graph of participants' ratings of their motivation on the day of testing (validation sample). 129

Figure 44.
A bar graph of participants' ratings of their general motivation to ensure that information regarding alcohol abuse is accurate (validation sample). 130

Figure 45.
A bar graph of participants' ratings of the test Web resource (validation sample). 132

Figure 46.
A histogram of participants' critical appraisal scores regarding the author of a Web resource (validation sample). 133

Figure 46.
A histogram of participants' critical appraisal scores regarding the content of a Web resource (validation sample). 134

Figure 48.
A histogram of participants' critical appraisal scores regarding the form of a Web resource (validation sample). 135

Figure 49.
Final re-specified structural model with regression weights 146

DECISION POINTS: DETERMINING FACTORS IN CRITICAL APPRAISAL OF HEALTH-RELATED WEB RESOURCES

The development of the World Wide Web (Web) has changed the way that people access information. This current shift is often referred to as the information age. Perhaps never before has access to information been as widespread or as easy. Part of the process of accessing and using information is deciding between different information sources and ensuring that the information meets ones needs. Regardless of where people obtain their information from, they must at some point, either explicitly or implicitly, pass judgment about the quality of it. (Hertzum, Anderson, Anderson, & Hanson, 2002). Not all decisions that people make, however, are valid.

With the increased access and ease with which people can post information on the Web, there has been a fear that inaccurate or misleading information is becoming prolific. . The factors that make the Web so appealing and that have led to its increasing importance in world, have also led to many of its supposed shortcomings (Mayer & Till, 1996).

For example, the easy, cheap unrestricted nature of the Web makes it possible for people to post information with very little cost, and with very little technical knowledge. This has allowed for unprecedented information sharing among individuals all over the world. However, because it is so cheap and easy, almost anyone can post information with very little in the way of expertise or funding. Consequently, there is little, if any, oversight necessary in order to post information, meaning that almost any site posted could potentially contain errors or misleading information (Metzger, 2005).

In addition, with the Web and the Internet making information searching easy for expert and layman alike, there are few barriers to access for anyone who wishes to find information and to try to make use of it. There is, then, a risk that users could find easily find and use inaccurate information.

Thus, the fact that there are few, if any, restrictions on what people can post, the fact that people can post with little technical knowledge, the fact that they can do so with very little cost, and the fact that there are relatively few barriers to anyone wishing to access information all combine to create a situation where it is understandable that some are concerned about the Web and how people are accessing and using information found there.

Research on this topic has come from a number of different fields (e.g., computer science, health, psychology as well as others). While each field has made significant and important contributions regarding how people find and evaluate information that they find on the Web, there is still no satisfactory framework to guide or situate this work. As the Web continues to become ubiquitous, it will become increasingly important understand the many different influences that lead people to make that decisions that they do. In order to establish this framework, I will first begin with a discussion of some of the important issues to consider when examining the critical appraisal of information from health-related Web resources.

It is important to emphasize that inaccurate information is not an issue that is restricted to the Web. Nor are the problems associated with how people read and make use of information specific to the Web. There are broader issues related to the problems and concerns that people have with the Web.

Information has always existed. The history of the development of universal, or near universal literacy in a number of ways parallels the current expansion of access to information made possible because of the Web. The development of literacy skills has been important regardless of the domain for decades, if not centuries. The increased influence of the web has simply highlighted this issue once again.

The Development of Literacy

The development of the ability to use written communication, has many probable causes and an extensive review is beyond the scope of this paper, however, a brief description of the issues that arose because of expanding literacy are informative with respect to the development of the Web will be provided.

The capacity for written communication has allowed societies to retain information and to pass that information on from generation to generation (Havelock, 1982; Olsen 1994). In addition, as literacy skills expanded, they allowed individuals to take on a much more active role in society. Having direct access to information allowed people to begin to make decisions for themselves without always having to rely upon authority figures for access to and interpretation of written documents.

A few of the factors that are hypothesized to have played a role in the expansion of literacy skills in the west are the development of the market economy, the reformation, and schooling (Venezky, 1991). In addition to these, advancements in technology such as the development of the printing press also allowed literacy to move beyond a privileged few, to many other members of society (Eisenstein, 1979).

Whatever the actual causes, literacy continued to expand and opened many doors that had previously been closed. Before the expansion of literacy, a few learned members of society, typically scribes and clerics, translated most texts. Eventually literacy spread to the nobility and then to the aristocracy but was still not a common skill among the general population. Thus, if one did have a question and needed an answer, one had to go to the holders of knowledge, namely, those few who were fortunate enough to read. Questions regarding theology, for example, were typically answered by an appeal to an authority figure (e.g., clerics) who would dispense the knowledge that he or she deemed appropriate.

With the advent of the printing press, written texts that were traditionally useful and available to only a few now began to be available to a much wider audience. At this time of expansion, there were attempts to restrict who could read and what it was that they were allowed to read. For example, Venezky (1991) claimed that in Europe there were restrictions of how many printing presses could operate and what types of publications they could print. In fact, people could even be executed for publishing books considered seditious (a practice not unheard of even in the 20th century). In the 15th century the Pope proclaimed that booksellers needed to get the permission of the church before creating a new book (Martin, 1988/1994).

Even as literacy was spreading, there were concerns about how people would interpret writing. In response to the concern that widespread literacy would become destabilizing for society, Clanchy (as cited in Olsen, 1994) claimed that the literacy of the general public would be restricted to the mechanics of reading and writing but that liberal education would still be preserved as the domain of a few elite. Thus having the ability to read was still not equated with the ability to understand and interpret what was being read. Even though texts were available, there were no guarantees that people could actually extract relevant information from the texts or make valid decisions regarding the accuracy of what they had found.

One emerging issue was that as literacy skills expanded and access to print material increased, the original holders of the knowledge began to lose control over how the uneducated masses understood and used the information they found. In addition to increasing lack of control, there were different interpretations that arose because of increased access to written documents. Before long there began to be a broadening of views as people started to question authority figures, to think for themselves, and to seek their own answers to question instead of relying on authority figures to explain to them what they needed to know (Martin, 1988/1994; Olsen, 2004).

There were still substantial costs associated with creating and maintaining the texts that held the information, however. So, although access to information was expanding, it was still limited due to the information seekers' abilities and the tendency, or need, to rely upon experts to disseminate information. Clerics, government officials, scholars and other experts continued to function as gatekeepers, replacing the previous gatekeepers, holding and dispensing information as they saw fit, a tradition that may continue in some respects today.

History of the Web

The Web is a recent creation; however, the idea of the Web is not necessarily a new one. The current formulation of the Web, although in a much more simplistic format than is now available, began in Europe in 1990 with the European Organization for Nuclear Research or CERN (http://info.cern.ch/).

Tim Berners-Lee originally conceived of the Web in Switzerland in 1990 (http://portal.unesco.org/ci/en/ev.php-URL_ID=12789&URL_DO=DO_TOPIC&URL_SECTION=201.html). Berners-Lee wanted to create a method for transferring information using the Internet. He developed a tool that connected hypertext and the Internet and allowed information to be shared among different computers. This idea was eventually expanded beyond the initial laboratory to create a Web where information could be shared between individual personal computers around the world.

The idea of a system of linking related files to allow for easy selection and cross-referencing was not new, however. Even as early as the 1940s, Bush (1945) described the concept for a tool that would allow for linking of information.

Given the technology of the time, Vaneveer Bush, the science advisor to the President of the United States, introduced the notion of connecting the information through what he termed a Memex (a desk that contained multiple microfiche screens). Information would be connected by topic and a user could view all of this related information at one time (Bush, 1945). Bush's Memex was a precursor to the development of hypertext, which, after a number of different iterations and developments, eventually led to the current form of the Web (Boechler, 2001; Conklin, 1987).

Since its development in the early 1990s, the Web has continued to develop and spread. Currently, users do not even need to use a personal computer to access the Web. Users can even gain access to the Web on devices as small and portable as cell phones. It is almost possible to locate any information, anywhere anytime. Although it is impossible to know exactly how many Web resources there are, as of October 2006 there were an estimated 97,932,447 Web sites (http://news.netcraft.com/), with more than 8 billion Web pages available.

As evidence of the ubiquitous nature of the Web, in 2004, 72% of all Canadians were Internet users while 90% of people between the ages of 18 and 24 reported using the Internet. Eighty-four percent of children were reported by their parents to use the Internet (Canadian Media Research Corporation, 2005).

Similar results have also been found in the United States. For example, Horigan (2006) in a survey of Americans found that 20% of survey respondents reported that they get most of their news and information about science from the Internet; this number was even higher for respondents under the age of 30. In this same study, Horigan noted that 73% of all Americans had internet access in early 2006, an increase from 63% in 2004.

As the Web and Internet continue to grow in scope and influence we are witnessing, perhaps, a similar process to the one that began as literacy skills were beginning to expand and improve. For example, some of the changes that are occurring now or have recently occurred seem to be breaking down barriers and removing traditional gatekeepers in favor of new, alternative approaches. The actual information was always there, there were simply more barriers in place blocking access to that information. Just as with the development of literacy skills, there was then and is now a certain level of literacy necessary, beyond the mechanics of reading, to access and use information appropriately.

In addition, just as there was then, there are attempts now to restrict the content of the Web and concerns about the information being published as well as people's understanding and use of this information. Recent articles describe some governments' attempts to control or restrict access to certain Web resources by their citizens (c.f., Goth, 2005). Interestingly, at the same time that attempts are being made to restrict access to some information, there are attempts being made to circumvent these restrictions. As in times past, restrictions do not always work as intended (http://psiphon.civisec.org/).

People are gaining greater access to information than ever before, and with that direct access comes greater individual responsibility to ensure that the information is accurate, appropriate, and worthwhile. If authority figures are bypassed, decisions regarding information quality need to be made by those accessing information. As in times past, simply having information available is not enough; information seekers and users need to have the ability to discriminate useful from non-useful information (Browne, Freeman, & Williamson, 2000).

As was demonstrated in previous centuries, restricting access or otherwise censoring information does not stop people indefinitely from accessing or finding the information if they want to find it. People need to learn how and understand why it is important to think critically about information that they find, whether that information is accessed via the Web or through some other source.

Critical Appraisal

Critical appraisal is "the process of systematically examining research evidence to assess its validity, results and relevance before using it to inform a decision (Hill & Spittlehouse, p. 1)." At the very least, critical appraisal of a Web resource should involve determining the following: (a) Who is the host or sponsor of the resource? (b) Who is the author of the information contained in the Web resource? (c) What is the purpose of the resource? (d) Who is the target audience? (e) How accurate is the content? (f) How objective is the coverage? (g) How comprehensive is the resource? (h) Is the resource current? and (i) Is the resource relevant to the needs of the user? (Varnhagen, 2002).

Studies describing how people search for and evaluate information generally show similar patterns, that is, that people tend to make very quick judgments about Web resources and that they appear reluctant to invest a lot of time in evaluating a resource if they are not sure that it is worthwhile. For example, studies by Huberman, Pirolli, Pitkow, and Lukose (1998) showed that people only move about two or three pages into a resource before looking for another. Similarly, Eveland and Dunwoody (1998) found that many information searchers left after viewing only one page and that most of those who stayed on a Web resource went to only two or three pages before they left.

Research by Bucklin and Sismeiro (2003) found that there is an inverse relationship between the number of pages viewed and the amount of time that users spend viewing a page. Users spend less time per page, the deeper into a Web resource they move. In addition, Bucklin and Sismeiro also found that the number of page requests on a given site was inversely related to the likelihood of future page requests, suggesting that users are less likely to return to sites that require moving deep into a site to access information.

In addition to a reluctance to invest time and effort in navigating deep into a resource to locate information, research has also shown that users do not always ensure that information is accurate. For example, in a study of college students' use of online resources, Metzger, Flanagin and Zwarun (2003) found that although students are increasingly using the Web for their schoolwork they are less likely to verify information that they find online.

Why do people make such seemingly rash decisions when it comes to evaluating information that they find on the Web? It could be either that they lack the ability to engage in critical appraisal or that they have the ability but lack the motivation to engage in such an effortful process. An additional perspective is that ability and motivational factors both combine to determine whether critical appraisal will occur.

In domains such as health, where using inaccurate information can have serious consequences, it is essential to understand how and why people choose certain resources over others. In conducting this research, my objective was to propose and validate a model of critical appraisal that examined the contributions that domain knowledge, knowledge and skill with the Web, knowledge of critical appraisal and motivation have on critical appraisal of a health-related Web resource. In addition, I examined whether measures of covariance between the indicators would lead to a more useful model. The purpose of this is to establish a framework that will guide and inform future research.

The Consequences of Using Information from the Web

With so many Web resources, and so much information available, one of the features of the Web that makes it so appealing on one hand, yet so frustrating on the other, is the fact that one can literally find anything. This makes the Web appealing because it is possible to locate information related to almost any topic. It makes it frustrating for information seekers because, as mentioned previously, there is little or no oversight. As a result, there are often no guarantees that information from a given Web resource is accurate. It is up to users to determine the usefulness and quality of any information that they find.

In many domains, the consequences of finding and believing bad information may not be life altering. For example, one needs only stop by a newsstand to sample many articles that are quite obviously fabricated. These fictional news articles, false though they are, have circulations in the millions. There is, then, obviously a market for this sort of information in the print media, and likewise, there are certain domains on the Web where it is not critical to have accurate information (e.g., resources designed purely for entertainment and other similar domains).

In many domains there are, however, real and significant consequences that can result from making decisions based on inaccurate or misleading information. Consider, for example, the domain of finance. Someone who does banking or shopping online, depends heavily upon the fact that the Web resource being viewed is, in fact, the Web resource of a reputable business, or at least that the Web resource contains accurate information and is secure (Metzger, 2006). When searching for investment advice, again, there is a certain amount of trust that one must place in the information from the Web resource being used. Failure to fulfill these requirements can have potentially devastating financial consequences. As a result, significant efforts have been put forth to ensure that those who are carrying out financial transactions on the Web can be assured of the security of the site they are using. In addition, there is also pressure from those who are using the Web to make sure that the sites are secure. Despite this fact, many malicious sites are constantly being set up to mimic reputable sites in order to steal money from users, a process referred to as phishing. Even with many precautions in place, users still need to exercise caution and discretion to ensure that they are making sound decisions.

Health Information and the Web

Within the domains such of health, the consequences of making decisions based on inaccurate information can be quite severe. The advent and subsequent popularization of the Web has led to unprecedented access to health-related information (Fox, 2006; Hansen, Derry, Resnick, & Richardson, 2003; Rideout 2001; Shuyler & Knight 2003). With that accessibility, however, comes an added personal burden of finding and evaluating Web resources and the information these resources contain (Eysenbach, Powell, Kuss, & Sa, 2002).

In the health field, physicians have been, and still are, the primary disseminators of information and knowledge. This is not to say that people could not access health information from other sources before the Web; in most cases, much of the appropriate knowledge was still there, typically located in libraries on university campuses. There was, however, a considerable amount of effort and expertise necessary to locate this information and use this information. One of the changes that has occurred in the past decade is that because of the Web, access to health information has now become much more widespread making it much easier for people to gain access to information regarding almost any topic (National Library of Medicine, 1998).

With the Web now allowing unfettered access to the same primary research and health information that health professionals are using, people no longer have to rely solely on these sources (e.g., , doctors), or gatekeepers for their information. As a result, there has been an increase in the number of people who are looking for and using health information from the Web. For example, when Medline was opened to the public in 1997, there was an increase in the number of searches of this medical database from 6 million in the year preceding the opening to 120 million in the year following the opening. This increase was largely attributed to the lay public (National

Library of Medicine, 1998).

Shuyler and Knight (2003) found that the percentage of people in the United States with Internet access who had used the Internet to search for health related information was 62%. In a related survey of adolescents, Hansen, Derry, Resnick, and Richardson (2003), found that more than 70% of adolescents had used the Internet to look for health related information. Rideout (2001) found that 68% of people between the ages of 15-24 had gotten health information online and that 24% had gotten "a lot" of health information from the Web.

In a more recent survey of how the American public use online health information, Fox (2006) found that 80% of people in the United States who had Internet access had used the Internet to search for health information and that the majority of these people had searched more than once. Fox also reports that, on average, eight million American adults will typically look for online health information every day. In addition, 58% of those who searched for health information said that the information they found had affected a health-related decision.

The Draw of the Web

The reasons why people are increasingly turning to the Web for access to health information are many and varied. Probably one of the most significant of these is the convenience that the Web provides. The ease with which people can access even the most sophisticated medical information has made the Web a popular venue, not just for the lay public, but also for professionals (Tang & Ng, 2006).

The government of Alberta, for example, has instituted a program to increase broadband Internet access to schools, libraries, hospitals, government buildings, post-secondary institutions, and potentially businesses and residences in remote, rural areas (http://www.education.gov.ab.ca/technology/SuperNet/). Through this program, infrastructure has been put in place to allow even rural and remote locations broadband Internet access.

The potential benefits of this are not only for the aforementioned institutions, but also for all residents of communities, as commercial broadband Internet connections can also be provided via this SuperNet. For those who live in remote areas, having access to databases of health-related information could, potentially, make it easier to make informed decisions regarding their own health or the health of others, especially when access to doctors and hospitals may be limited.

In addition to ease of access, information on the Web is available on a more-or-less self-serve basis. This may be desirable for many people who want to know whether they should be concerned about a health related problem. Or, it could also be useful for those who want to get information about a particular health-related issue but do not want to ask someone else, or to let others know of their concern.

Even in urban centers, with ready access to health care, there may still be pressure for many people to make decisions regarding the need for professional care or for a diagnosis before they go to see a health care professional. That is to say, people who rush to see a health-care professional when there is nothing actually wrong with them take up space that could be used for someone who really does need help. Conversely, the Web could provide information that could lead one to seek professional help, thus avoiding more serious problems or complications that could result from delays in acquiring help. Access to health information on the Web can help to inform people regarding their own health.

The Web also allows a certain level of anonymity that is not possible with other methods of access. For example, entire online communities have developed for health-related issues that carry a level of social stigma. One of the most well developed online health communities is for those suffering from eating disorders (Something-fishy.org), but many others have also been developed for many different maladies and concerns (e.g., alcohol abuse, self-mutilation, suicide, depression, etc.). The Web allows people a venue to gather information to determine whether they have a problem, find help, contact others, find support, and many other things that they may not feel comfortable doing face to face.

There is also the cost saving potential that the Web has for people and for society. It costs money to consult professionals. Any time that one can access information without having to consult a professional, this will result in potential cost savings for the user. Even in societies with publicly funded medical systems, there is a cost associated with consulting professionals. The Web has the potential to allow people to deal with minor issues on their own without having to constantly ask professionals.

Assessing the Quality of Online Health Information

With an increasing number of people accessing information and more and more health-related Web resource emerging, the task of sifting through the available information becomes a daunting task. This is not a problem as long as people make informed decision about the information they find. If people cannot or do not make informed decisions then they may be making important, health-related decisions on information that is not correct, or potentially even misleading.

As access to information has become more commonplace, there has been increasing emphasis on training both undergraduates, and professionals, medical and otherwise, about the importance of critical appraisal.

In contrast to the lay-public, physicians typically have a number of different tools available to them when it comes to evaluating health information (i.e., domain expertise and extensive educational practice with evaluation of information). It is not clear, however, from the current research whether physicians actually use these tools to critically evaluate all health-related information they find on the Web.

Despite the fact that physicians and other experts are, perhaps, better able to decide between quality information and poor information, it is the lay public that often has to decide whether information is good or not. The problem, then, is not whether there is too little information, or whether there is too much information, the problem seems to be that people are not using any systematic approach when determining whether to believe information or not. That is to say, people, in general, are not engaging in critical appraisal even though there may be important reasons why they should.

As research by Eysenbach and Köhler (2002) has demonstrated, the top two strategies reported by users to assess the credibility of health Web resources are that they try to find information about the source of the site (i.e., who created it) and that they looked to see that the design of the site looked professional. While it is encouraging that people report that they are looking for information about the source, it is less encouraging that they also use the design of the Web resource as an indication of its usefulness and believability. In addition, and perhaps even more alarming, even though people reported that they looked for information about the source, they seldom report checking the "about us" section available on most sites. When asked, most users could not even remember where the information came from. Flanigan & Metzger (2007) found that those who claimed that they verified information were actually less likely to do so. Fox (2006) found that three quarters of those who seek heath-related information online do not check the source and date of information that they find.

Factors Influencing Critical Appraisal

One of the reasons why people do not engage in critical appraisal is that there may simply be too much information for them to process and that they become overwhelmed by the sheer volume and begin to experience cognitive overload (Niederhauser et al., 2000). While it might make some intuitive sense that people are not engaging in critical appraisal because they are overwhelmed by the sheer volume of information that is available, this approach does not adequately capture the complexity that is the Web. Human processors are by nature relatively efficient. For example, even though there is constant auditory stimulation, people do not tend to become overwhelmed by this constant barrage of competing stimulation. Instead, what tends to happen is that people become selective in terms of what information they process (Chen & Chaiken, 1999; Navon & Gopher, 1979). There is no reason to expect that the process would be any different for evaluation of information from the Web.

In examining critical appraisal, then, it is important to consider all of the factors, both ability and motivational, that may contribute to the likelihood that critical appraisal will take place. As can be seen in Figure 1, the potential contributing factors that I examined are domain knowledge, knowledge and skill with the Web, knowledge of critical appraisal and motivation.

Domain Knowledge

One reason why people may not be critically appraising Web resources could be lack of domain knowledge. As with expertise with critical appraisal, expert domain knowledge equips information searchers with additional cognitive resources (Jenkins, Corritore, & Wiedenbeck, 2003). Research from cognitive psychology has shown that experts are generally able to perform better than non-experts on domain specific tasks in a variety of different domains (c.f., Barfield, 1986; Chase & Simon, 1973; Vicente, 1992).

Chase and Simon (1973) proposed a theory to explain how experts' domain knowledge allows them to remember more. Chase and Simon examined how expert

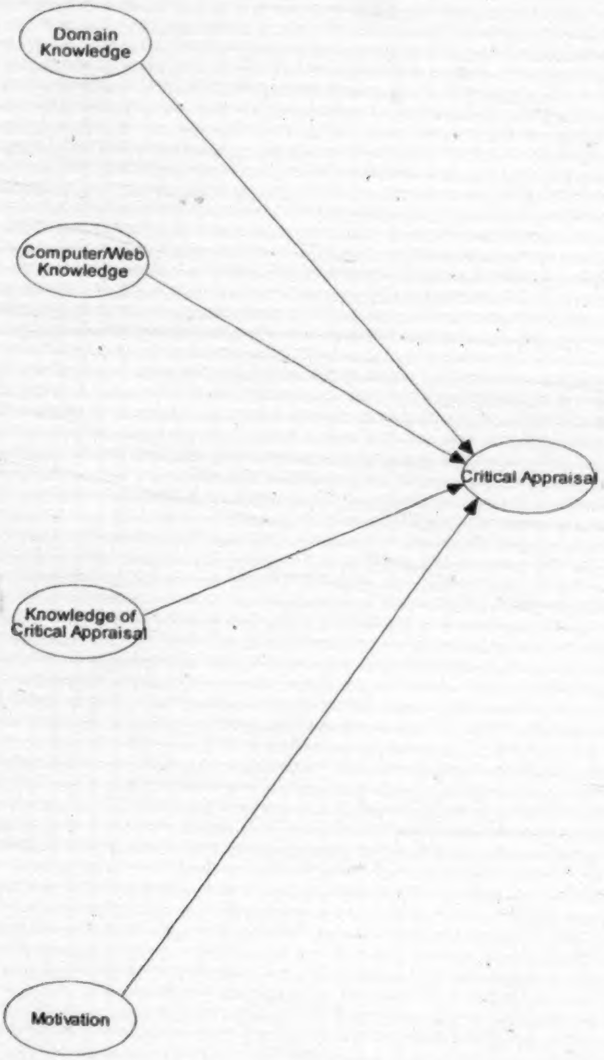
Domain Knowledge
Computer/Web Knowledge
Critical Appraisal
Knowledge of Critical Appraisal
Motivation

Figure 1. A proposed model of the factors related to critical appraisal of health related Web resources.

chess players compared to novice chess players when asked to remember various positions of chess pieces. According to Chase and Simon, experts were able to remember the positions of chess pieces better than non-experts on the same task because of the experts' pre-existing schemas and because they could more effectively chunk information. In addition to being able to remember more task related information, expert chess players were also able to retain this information with a much shorter exposure to the chess board than the non-experts were. Because of the existing schemas that experts have formed for the domain in question, they do not need to create new schemas when presented with new material. They are much better able to integrate new information within their existing cognitive schemas.

More effective use of cognitive schemas may also be useful for searchers and evaluators of information from the Web. As evaluators come to new Web resources, they can more efficiently recognize and process evidence regarding information quality and site credibility.

Vicente and Wang (1998) also proposed an explanation for the benefit that experts gain in memory processing called the constraint attunement hypothesis. According to this hypothesis, domain experts can outperform non-experts on tasks because they are more attuned to the goal-relevant constraints. Thus when experts and non-experts perform a task, the more constraints there are, the better the experts will perform relative to the non-experts. The benefit that the experts gains, according to this hypothesis, is that they have an extra toolset that helps them to deal with any complications or problems that arise.

In addition, knowledge of the prominent terms, names, and other specifics of a particular domain may allow a searcher to more effectively critically appraise a Web resource (Alexander, Kulikowich, & Schulze, 1994). The absence of this knowledge may leave a searcher with fewer options. For example, searchers may know that it is important to determine who the author of a particular resource is. However, even if they find this information, if they do not know the prominent names in the field they may not have enough information to make an informed decision. Domain novices lack the relevant schema to be able to situate themselves properly within the domain (Cole & Leid, 2003). The potential results of this are confusion, and possibly selection or use of substandard Web resources and material. As Bailin, Case, Coombs and Daniels (1999) state, "…the depth of knowledge, understanding and experience persons have in a particular area of study or practice is a significant determinant of the degree to which they are capable of thinking critically in that area" (p. 290).

In addition to the ability to critically appraise information from the Web, domain experts may have a higher level of literacy related to their specific domain. Even if novice searchers are able to find the information that they are looking for, they may not be able to understand what they are reading. D'Alessandro, Kingsley and Johnson-West (2001), for example, in a study of Web resources designed for pediatric patient education found that most of the information designed for patients to read is not even written at a level that would allow for the average adult to read and comprehend. Similar results have been found by researchers examining many different kinds of health information located on the Web (c.f., Graber, Roller &

Kaebler, 1999).

Knowledge and Skill with the Web.

In addition to, or instead of, being a domain expert, people may also have more or less Web and computer knowledge (Hodkinson, Kiel, & McColl-Kennedy, 2000). The ability to maneuver skillfully through the Web could be incredibly valuable. Consider for example, someone high in expertise at critical appraisal and high in domain knowledge but low in expertise with the Web. Such a person would, perhaps, be less skilled at finding those things that are important for critically appraising a Web resource, or have less of an understanding about basic navigational issues (Eveland & Dunwoody, 2000; Hargittai, 2002). Would he or she know how to parse a URL or where to look to find information about when the resource was created or updated? Hargittai claims that that those who are lacking in knowledge of how to use the Web have a hard time with even the most basic of tasks. If searchers are devoting a significant portion of their available cognitive resources just to navigate through the Web, this might leave even fewer cognitive resources for the task of critical appraisal (Niederhauser et al., 2000). In addition, research on decision making in naturalistic settings has found that experts are better able to use their prior knowledge in new situations which allows them to make more effective decisions (Randel & Pugh, 1996).

Knowledge of Critical Appraisal.

Fogg et al. (2002) examined how people determined the credibility of online resources. What they found, while not surprising, is somewhat disheartening for any who consider critical appraisal important. The majority of people in their study did not use “rigorous criteria” (p. 6), but instead based their credibility judgments on the look of the site. Why would the average person be so adverse to critical appraisal? One possible explanation could be the lack of knowledge of the average person about critical appraisal.

This is supported by findings from Stanford, Tauber, Fogg, & Marable (2002) who found that experts were more likely than novices to consider factors other than the design look in their appraisal of the credibility of a Web resource. Possibly, experts differ from novices because experts know what things they should be looking for in order to determine credibility of a Web resource. Taylor et al. (2000,) in a review of critical appraisal training programs for physicians, found an overall improvement of 68% after the training program suggesting that knowledge of critical appraisal does make a significant difference in people's abilities to critically appraise. In addition, Taylor et al. suggested that unless given specific training, people may not spontaneously critically appraise.

Motivation.

Is it enough, however, to have the knowledge necessary to critically appraise a Web resource? In addition to factors related to people's abilities in critically appraising Web resources, motivational factors are important to consider (Sawasdichai & Poggenpohl, 2001). While we sometimes believe that people are by nature “optimizers”, that is, that people continually try to maximize benefits, perhaps this is not always the case. In the field of economic theory Simon (1957) postulated that organisms adapt well enough to get by and that they do not normally optimize. This point is central to Simon's bounded rationality principle. This principle states that organisms generally act in a manner that is nearly optimal, but that this behavior is often bounded by conditions. According to the bounded rationality principle, expertise and knowledge could be potential bounding conditions, but there may be others as well. Thus, based on their circumstances, users may be operating nearly optimally but the bounding conditions, of which others may not be aware, may make it look like they are operating at a sub-optimal level.

Petty and Cacioppo (1986), in their research on persuasion and attitude formation, proposed a model which offers a clue as to what some of these other bounding conditions might be. According Petty and Cacioppo's Elaboration Likelihood Model (ELM), people want to have correct attitudes and they want to base their attitudes on correct information, but there are things that can affect their willingness or motivation to actually do this. Ability factors such as domain and technical expertise, message comprehensibility, distraction, etc. are important determinants of whether people will critically appraise. However, it is motivational factors which determine whether ability factors will be important or not. If evaluators are not motivated to process information they are less likely to critically appraise regardless of their abilities to do so (Claypool, Mackie, Garcia-Marques, McIntosh, & Udall, 2004; Petty & Wegener, 1999; Rothman & Schwartz, 1998).

While there is little debate that motivation is an important factor in determining whether people will engage in effortful processing, motivation is, itself, a difficult term to operationalize and could possibly include many different influences. Two motivational factors that Petty and Cacioppo (1986) claim are important to whether people will be likely to critically appraise are (a) personal relevance and (b) personal responsibility.

Personal relevance according to Petty and Cacioppo (1986) occurs when an issue has, or is expected to have, some personal consequences for the evaluators. Increased personal relevance has been shown to increase people's willingness to evaluate information more carefully (Claypool et al. 2004). For example, a parent researching treatment options for her child recently diagnosed with autism spectrum disorder may be more motivated to critically appraise a Web resource than would a student assigned to research the same topic.

In addition to personal relevance, people may be more motivated to critically appraise if they feel some pressure to ensure that their evaluations are valid. Petty and Cacioppo (1986) refer to this as personal responsibility; however this same phenomenon has also been referred to as fear of invalidation (Fazio & Towles-Schwen, 1999).

Personal responsibility or fear of invalidation has been shown to be an important factor in increasing the likelihood that people will critically appraise (Sanbonmatsu & Fazio, 1990; Schuette & Fazio, 1995). Schuette and Fazio, for example, demonstrated that when fear of invalidation is increased, in their case by informing participants that their evaluations would be compared to a standard and that they would be questioned regarding their evaluations, participants were more likely to use a cognitively demanding approach.

According to Petty and Cacioppo (1986), as people feel more responsibility for evaluating information, they are more willing to critically evaluate the arguments. Whether someone experiences this has typically been inferred through behavior, or by asking participants to report the level of relevance they feel the topic has for them and the level of responsibility they feel in ensuring that the information is accurate.

While there may also be personality factors that are important in determining whether someone will critically appraise or not, in general, critical appraisal seems to be a reflection of more ephemeral factors. Given the right circumstances, even the best critical evaluator may choose not to critically evaluate.

A Modified Approach

In order to understand how ability and motivation factors interact and influence each other, I propose a modified model of critical appraisal of health-related Web resources based on ELM (Petty and Cacioppo, 1986). As with ELM, this modified model will examine both the effects of ability and motivation. Unlike ELM, however, the proposed model is non-sequential. The factors believed to be important in determining whether critical appraisal will occur are knowledge or skill with critical appraisal, domain knowledge, knowledge or skill with the Web, and motivation to critically appraise. In contrast to other models of persuasion and critical appraisal that portray critical appraisal as a sequential process, Figure 2 illustrates the possible non-sequential interactions between these factors.

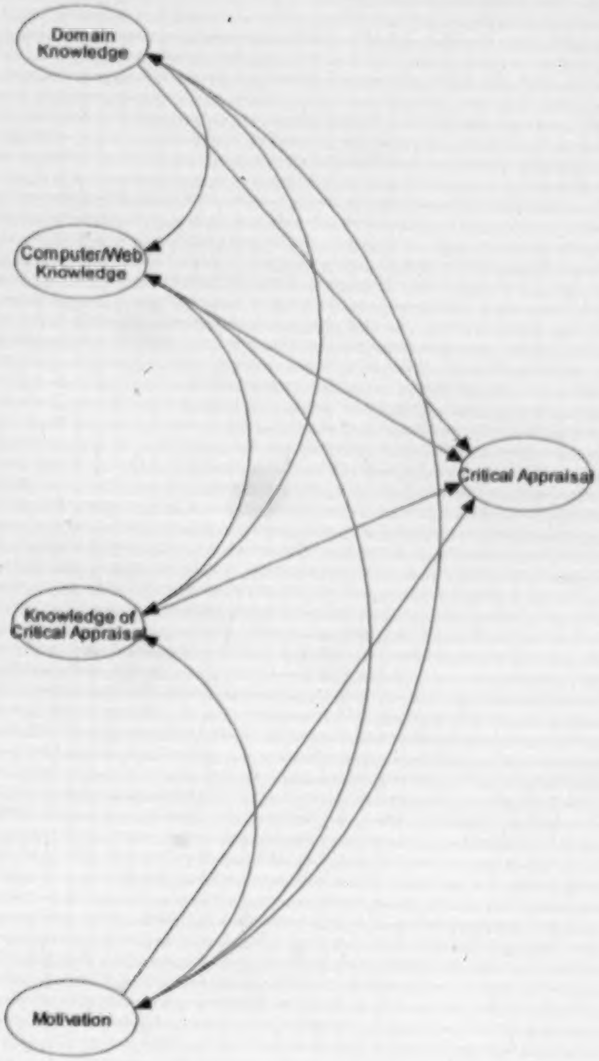
Domain Knowledge
Computer/Web Knowledge
Critical Appraisal
Knowledge of Critical Appraisal
Motivation

Figure 2. A model of the possible interactions between factors related to critical appraisal of health related Web resources and critical appraisal.

As can be seen, all of these factors could possibly be interrelated. In addition, there are no a-priori assumptions of temporal precedence. That is to say, to begin with, no factor precedes any other factor in importance or influence; rather, the assumption is that all the factors could be influencing critical appraisal simultaneously as well as influencing each other.

Domain knowledge is the amount of domain specific knowledge that a person has. The goal of including this measure is to quantify the amount of knowledge that participants have. A person with high domain knowledge may apply this knowledge when evaluating Web based resources by using this knowledge to effectively evaluate the quality of information and the credentials of the author.

In addition, domain knowledge might allow searchers to develop better formulated search strategies and more appropriate evaluation criteria. Lack of domain knowledge may hinder evaluation because there is no base of reference. For example, searchers may find information that is necessary for critical appraisal but because of their limited domain knowledge, the usefulness of the information may be limited. All other things being equal, it would be expected that people high in domain knowledge would outperform those who score low in domain knowledge (Alexander et al., 1994; Hassebrock, Johnson, Bullemer, Fox, & Moller, 1993; Vincente & Wang, 1998).

Knowledge and Skill with the Web refers to the ability to search and navigate the Web. People low in Web skill might be expected to have less efficient and effective searches. Again, the purpose of including this measure is to quantify knowledge and skill with the Web. If it takes all, or most, available cognitive resources, it is possible that there will be fewer available resources that can be effectively applied to evaluation. In addition to this, skill with the Web might allow for faster, more effective search strategies that will reduce the amount of cognitive resources that need be applied to finding the information (Jenkins et al., 2003).

Another reason why knowledge and skill with the Web is important is that increased knowledge may allow an information evaluator to find critical information much more easily (Randel & Pugh, 1996). For example, if an evaluator would like to determine whether the author of information on a Web resource is credible, a skilled Web user may be able to find this information even when it is not apparent. Evaluators scoring high on knowledge and skill would be expected to outperform those scoring low (Hargittai, 2002).

Knowledge of Critical Appraisal is the amount of knowledge that people have regarding the process of critical appraisal. Being a skilled critical appraiser would be beneficial for critical appraisal because if people know what is important to know when critically appraising, they will be more likely to actually seek out this information. For example, knowing that it is important to know who sponsors the Web resource may make it more likely that an evaluator will actually look for or pay attention to this information on the Web resource. In contrast, however, if a person has very little knowledge of critical appraisal, he or she may not find, notice, or even understand the significance of relevant information. Having knowledge of critical appraisal should increase the likelihood of critical appraisal occurring (Taylor et al., 2000).

Motivation is that aspect of an evaluator that increases the likelihood that the evaluator will expend the cognitive energy necessary to critically evaluate. Typically motivation is generally inferred post-hoc by observing whether participants actually critically appraise. In order to break from circularity of this argument, it is important to determine whether participants are motivated prior to the task. Two measurable, factors that have been shown to influence motivation are personal relevance and personal responsibility (Claypool et al. 2004; Sanbonmatsu & Fazio, 1990; Schuette & Fazio, 1995).

As people feel that an issue has increasing relevance for them they will be more likely to expend the energy to effortfully appraise a Web resource. Issues or topics that are not personally relevant will, in general, not increase the likelihood that critical appraisal will take place.

Similarly, as people feel that they are personally responsible for the accuracy of information (e.g., if they were going to be required to justify their answers) they are more likely to critically evaluate (Schuette & Fazio, 1995). When taken together, high personal relevance, high personal responsibility should lead to an increase in the likelihood that critical appraisal will occur.

Critical Appraisal is assumed to have occurred if an evaluator correctly accepts strong arguments and correctly rejects poor arguments. The proposed model is an attempt to explain what role the aforementioned contributing factors play and possibly to predict what conditions will lead to increased likelihood critical appraisal.

So, in addition to measures of domain knowledge, knowledge and skill with the Web, knowledge of critical appraisal, and motivation, a measure of critical appraisal was taken (i.e., whether or not critical appraisal actually occurred). The measure of critical appraisal included both a global assessment of the Web resource and a retrospective report of different measures that are assumed to be involved in critical appraisal.

When taken in isolation, each of the factors (i.e., domain knowledge, knowledge and skill with the Web, knowledge of critical appraisal, and motivation) has predictable effects on the likelihood of critical appraisal. In reality, however, the factors are probably not orthogonal. One key aspects of this model is that it does not assume, a priori, that factors related to critical appraisal contribute to critical appraisal in isolation, nor that they contribute in any predetermined linear or sequential fashion. At any given time one or more than one of the factors mentioned can influence critical appraisal directly while also influencing other factors or moderating their influence on critical appraisal.

Thus two major hypotheses in this dissertation are that the covariance measures between each of the factor are important to the overall strength of the model and that motivation has not only a direct impact on critical appraisal, but also serves to moderate the influences of the other ability factors. While traditional hypothesis testing might consider it a weakness that no specific relationships are hypothesized a priori, this is an important feature of the process of creating a valid, externally valid model. Because it is not known what effect each of the factors has when considered in light of all of the others, any speculation about the possible strength of relationships is premature.

Although critical appraisal is more a process than a static event, this study captures only a snapshot of evaluators' behaviors. While this may initially appear to limit the usefulness of the model, it is still an important first step in determining what factors are important for critical appraisal and will form a foundation upon which future research may be based.

Structural Equation Modeling

To capture the dynamic nature of the critical appraisal process, I used structural equation modeling (SEM) (Hayduk, Mah & Carter-Snell, 2002). All analyses were conducted using AMOS 4 (Arbuckle, 1994). Structural equation modeling is an effective statistical technique for model testing such as this. As the first step in this process involves establishing a baseline model from a number of observed and latent variables, a method that allows for confirmatory factor analysis will enable me to create and test the proposed model. In addition, a method that allows for post-hoc modification of the model in a more exploratory fashion will allow for a more parsimonious, and useful model.

One strength of SEM is that it allows for the simultaneous comparison of contributing factors. This is critical for this model because critical appraisal may not be a linear or stepwise process. Each factor involved in the process may be modified and could, potentially modify every other factor. In addition, with respect to the temporal sequence of events leading to critical appraisal, the hypothesized factors are possibly influencing each other simultaneously rather than sequentially as some have speculated.

SEM is a powerful tool, yet it has drawn some criticism as it is also referred to causal modeling. While this aspect has been controversial and there are those who would claim that SEM should not be used because of these supposed shortcomings, there is still some utility that can be gained from SEM even without being drawn into the debate over causality. In this study, there are no claims of causality; rather, SEM is being used to determine whether a plausible model can be found to explain, in part, the process of critical appraisal of health related Web resources. Validation with SEM does not prove that this is the way that critical appraisal takes place, but it does suggest that this may be one possibility.

To study the factors that are important for determining whether critical appraisal will occur, I had psychology students examine and rate one of two health-related Web resources. One of the challenges of performing research on critical appraisal is that because the research setting is an artificial setting, it can be hard to actually ensure that the research is externally valid. I chose two different domains that I believed would be relevant to undergraduate students, and thus would, hopefully, lead them to act in a manner that was similar to what they would do if they were choosing to search for information on their own.

The two topics that I chose were alcohol abuse and eating disorders. While I did not expect that every student would have a personal interest in these topics, they are widespread enough in undergraduate communities that there was a good chance that most people had some familiarity with the issues and that some knew a lot through personal experience or by having an acquaintance who had experience.

Before having students examine one of the health-related Web resources, I had them fill out a number of surveys to measure their domain knowledge, their knowledge and skill with computers and the Web and their knowledge of critical appraisal. In addition, I also had them answer questions regarding their own motivation to search for information and to ensure that it was accurate.

STUDY 1: ESTABLISHMENT AND RE-SPECIFICATION OF THE BASELINE MODEL

Methods

Participants

The baseline sample consisted of 210 undergraduate students in introductory psychology classes recruited through a research participant pool. Participants received partial course credit for their participation in this study. Of the 210 participants, 133 were female (approximately 63%) and 77 were male (approximately 37%). The mean age of participants was 19.3 years, *SD = 3.15*.

Measures and Scoring

Domain Knowledge. Appendix A contains test of domain knowledge. Questions on this test tested participants' knowledge of alcohol abuse as well as eating disorders. Topics tested covered a broad range of issues related to alcohol abuse (question 3-17 in Appendix A) and eating disorders (questions 18-42 in Appendix A). Scores on these tests of knowledge could range from 0 to 15. In addition, this test asked participants to self-report their level of domain knowledge related to alcohol abuse (question 1 in Appendix A). Self reported domain knowledge was measured using a 5-point likert-type scale. Responses from the two questions were summed resulting in scores ranging from 1 (indicating little to no domain knowledge), to 5 (indicating a high level of domain knowledge). Cronbach's Alpha for domain knowledge was .52. Appendix B contains an answer key to the questions that were asked.

Knowledge and Skill with the Web. Appendix C contains a measure of knowledge and skill with the Web. Participants answered using 5-point likert-type scales and answers were summed with categories; higher scores indicating a higher level of knowledge and skill with the Web. This test was used to assess participants' knowledge of the Web and participants were asked to self-evaluate how much they liked working with computers and the Web (questions 2 & 3 in Appendix C) how frequently they used the computers and the Web (questions 4, 5, & 6 in Appendix C), and their knowledge and skill computers and the Web (questions 7 & 8 in Appendix C). Cronbach's alpha for knowledge and skill with the Web was .77.

Knowledge of Critical Appraisal. Appendix D contains a test of knowledge of critical appraisal. Knowledge of critical appraisal was measured by having participants report strategies that they were aware of that they could lead to critical appraisal. These strategies were rated as valid or not valid based upon criteria set by Varnhagen (2002). All answers were coded by research assistants. Initial inter-rater reliability was high (0.90) and all discrepancies were resolved through discussion. Valid strategies were assigned a one and invalid strategies were assigned a zero. These strategies were then multiplied by the reversed-scored, reported frequency. This resulted in scores ranging from 0 (for invalid strategies) to 4 (valid strategies that are used all of the time) for each strategy reported. This measure contains a number of sub-measures such as knowledge of the author (question 43 in Appendix D), the credibility or believability of the author (question 44 in Appendix D), the purpose (question 45 in Appendix D), the accuracy (question 46 in Appendix D), the objectivity (question 47 in Appendix D), the comprehensiveness (questions 48 in Appendix D), and currency (question 49 in Appendix D). Cronbach's alpha for knowledge of critical appraisal was .52.

Motivation. Appendix E contains a survey designed to determine the participants' level of motivation at the time of the evaluation. This test contains items designed to measure personal responsibility (questions 27, 28, 29, 30, 31, 32, 33a, & 33b in Appendix E) and personal relevance (questions 33c, 33d, 33e, 34, 35, 36, & 37 in Appendix E) with respect to alcohol abuse. For both personal responsibility and personal relevance, scores were created by summing across answers for each category (higher scores indicating higher levels of motivation and lower scores indicating low motivation). In addition to these measures, an overall motivation measure (question 39 in Appendix E), and a current motivation level (questions 38 in Appendix E) were determined through a self-evaluation using a 5-point likert-type scale. Cronbach's alpha for motivation was .64.

Critical Appraisal. There are 5 different measures of critical appraisal. The first is an evaluation of an actual Web resource related to alcohol abuse found at http://familydoctor.org. Participants evaluated this Web resource using a 5-point likert-type scale (question 1 in Appendix G). In addition to the evaluation of the Web resource, participants filled out a survey (questions 2-40 in Appendix G) of factors related to critical appraisal. These were items that were most frequently deemed to be important for critical appraisal based on a review of the relevant literature. Participants were asked to report, retrospectively, what information they used to evaluate the Web resource. Critical appraisal was measured by comparing participant reports of their own behavior to pre-determined assessments of whether questions were important for critical appraisal. Most of the items used would be important for critical appraisal, however, there were some items that, while important to one's overall opinion of a site, were not directly related to critical appraisal (items 12-15, 31-33, 35, 36, 39, & 40).

In order to capture the fact that critical appraisal involves not only knowing what to pay attention to but also what to ignore, these items were scored differently. For example, for those items that were pre-determined to be important for critical appraisal, participants scored a 1 when they reported that they had looked for this information the pre-determined evaluation and a 0 when they reported that they did not look for this information either because they did not think that it was important or because they did not think to look for that information.. If the item was pre-determined to be unimportant for critical appraisal, participants scored a 1 if they either said that they did not look for this information because it was not important or they said that they did not think to look for this information. They scored a 0 if they thought that this was important information to know.

Each separate factor was used as separate measure of a component of critical appraisal. The determination of whether or not participants engaged in critical appraisal was made by examining participant responses regarding the author (questions 2-15 in Appendix G), the purpose (questions 16 & 17 in Appendix G), the content (questions 18-30 in Appendix G), and form (questions 31-40 in Appendix G)

Procedure

Testing occurred in two sessions for each participant. The second session was conducted no sooner that 1 week following the first session. As it was essential to get an accurate measure of critical appraisal, the lag time between sessions was to ensure that there was no influence of training on subsequent testing sessions. Participants were tested in groups of approximately 30 in the first session and 12 in the second session. The discrepancy in group sizes between sessions was because each participant needed to have a computer in the second session. The testing facility had 13 computers available for participants to use.

In the first testing session, participants were asked to complete the measures for domain knowledge, knowledge of computers and the Web, and knowledge of critical appraisal. These were all presented in paper format and participants had a maximum of one hour to complete these measure. At the end of the first session, participants were invited back for a second session.

During the second session, participants were randomly assigned to individual testing rooms that contained computers. The Web resource was loaded onto each of the computers, and then the monitor was turned off.

When participants entered the testing room, they were instructed to complete the motivation survey (Appendix E) which was face down on the desk. After completing this survey, participants were instructed to turn the monitor on and to examine the resource that was loaded onto the computer. Participants were not given time limits, but were instructed to take as much time as they felt was necessary for them to get a "good feel" for the site. When participants finished viewing the resource, they were asked to leave the testing room (so that they could not refer to the Web resource while completing the critical appraisal measure) and were given the critical appraisal measure (Appendix G).Participants were instructed that they should first rate the site and they they were asked to fill out the critical appraisal measure by reporting those things that they looked for in the site in order to be able to make a judgment about the site. This was a retrospective reporting of the types of information upon which they based their evaluations of the site. After completing this measure, the testing session ended and participants were dismissed.

Results

Specification of Baseline Model

Descriptive Statistics

Domain knowledge. Participants were able to answer correctly slightly more than half of the 15 domain related question having to do with alcohol abuse (M = 8.07, SD = 2.23). Self-reported domain knowledge was 3.05 (out of 5) with a standard deviation of 0.91. Both distributions were approximately normal (see Figures 3 and 4 respectively).

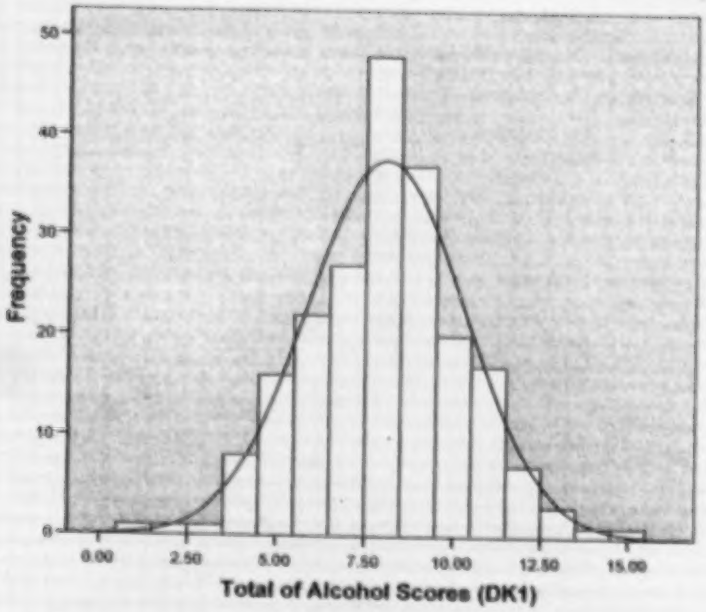

Figure 3. A histogram of the frequency of test of alcohol abuse domain knowledge (baseline sample) with normal curve.

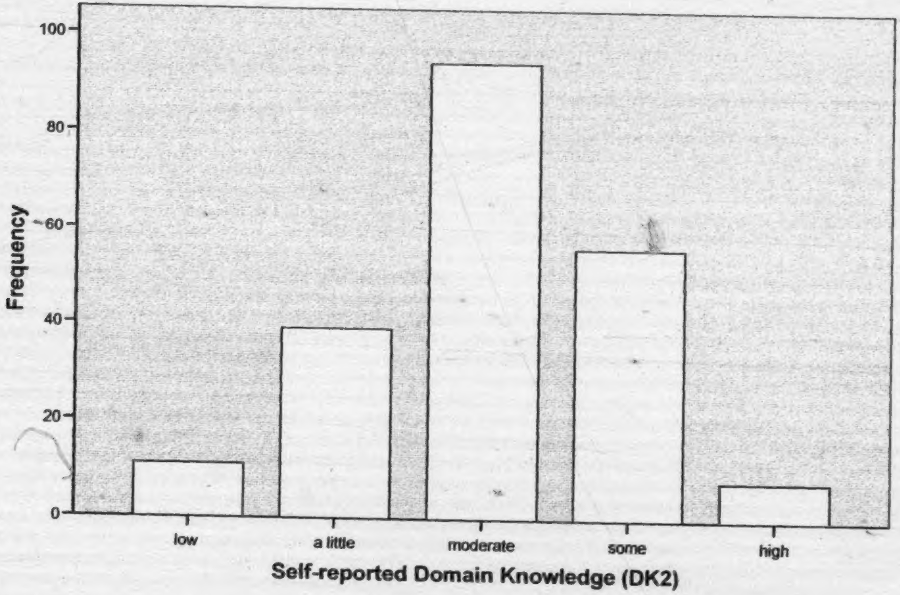

Figure 4. A bar graph of the frequency of self-reported domain knowledge (baseline sample).

Knowledge and skill with the Web. The mean response for how much participants liked computers and Web (CK1) was 7.88 (out of 10) with a standard deviation of 1.47. The resulting distribution was slightly skewed to the left and was truncated at the top end of the scale see Figure 5).

Questions regarding participants' frequency of use of the computer and the Web (CK2) resulted in a nearly normal distribution (see Figure 6). Mean frequency reported was 10.60 (out of 15) with a standard deviation of 1.68.

The distribution resulting from questions regarding skill with computers/Web (CK3) was nearly bi-modal (see Figure 7) with a mean of 7.40 and a standard deviation of 1.52.

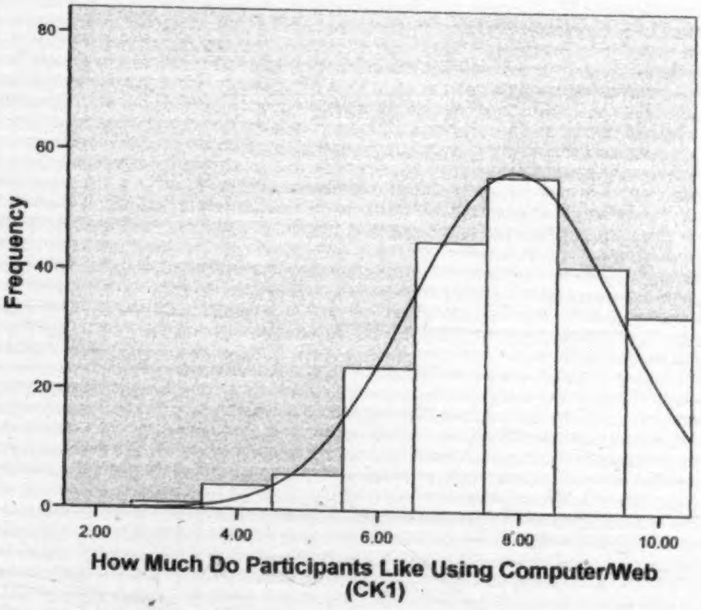

Figure 5. A histogram of the frequency of participants' ratings of how much they like to use Computers and the Web (baseline sample) with normal curve.

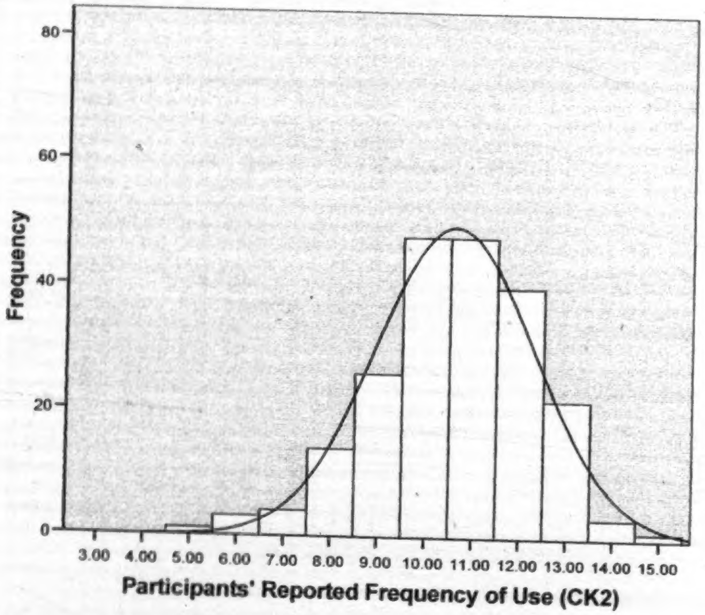

Figure 6. A histogram of the frequency of participants' ratings of how frequently they use Computers and the Web (baseline sample) with normal curve.

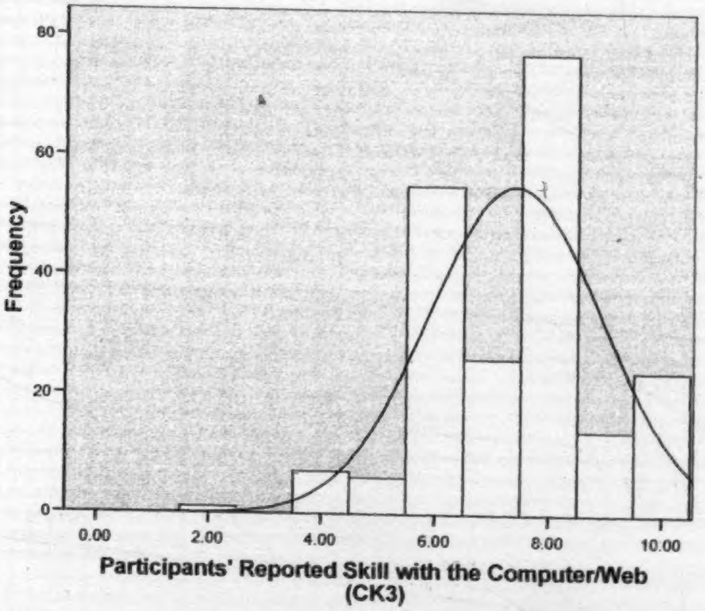

Figure 7. A histogram of the frequency of participants' reported skill with Computers and the Web (baseline sample) with normal curve.

Knowledge of critical appraisal. Knowledge of critical appraisal was measured using several different indicators. Each measure had a minimum score of 0 for no valid strategies reported and 4 for each valid strategy reported (higher scores indicating more valid strategies used more frequently). Knowledge of the author (KA1) had a mean of 2.25 and a standard deviation of 1.81 (see Figure 8). Strategies for assessing the credibility/believability of the author (KA2) had a mean of 0.81 and a standard deviation of 1.39 (see Figure 9). Strategies for assessing the purpose of the Web resource (KA3) had a mean of 1.48 and a standard deviation of 1.41 (See Figure 10). Strategies for assessing the accuracy of the Web resource (KA4) had a mean of 0.85 and a standard deviation of 1.15 (see Figure 11). Strategies for assessing the objectivity of a Web resource (KA5) had a mean of 1.39 and a standard deviation of 1.62 (see Figure 12). Strategies for assessing the comprehensiveness of a Web resource (KA6) had a mean of 0.74 and a standard deviation of 1.25 (see Figure 13). Strategies for assessing the currency of information found on a Web resource (KA7) had a mean of 1.50 and a standard deviation of 1.16 (see Figure 14).

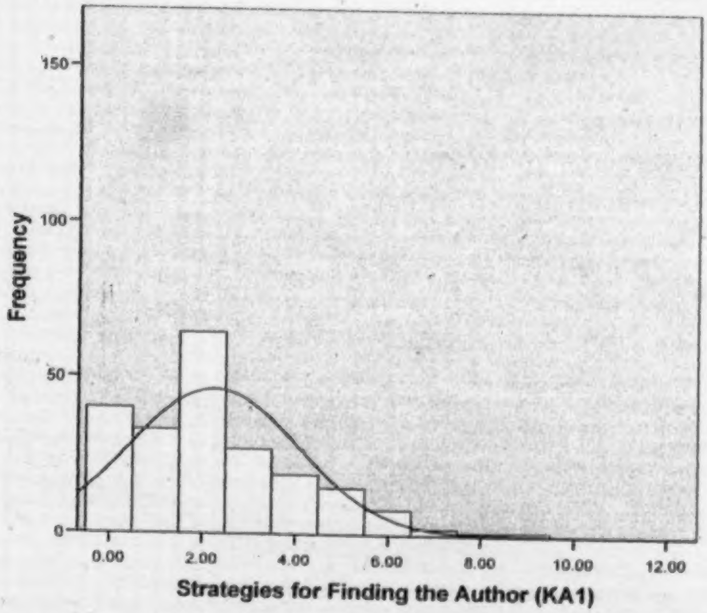

Figure 8. A histogram of participants' strategies for finding the Author of a Web resource (baseline sample) with normal curve.

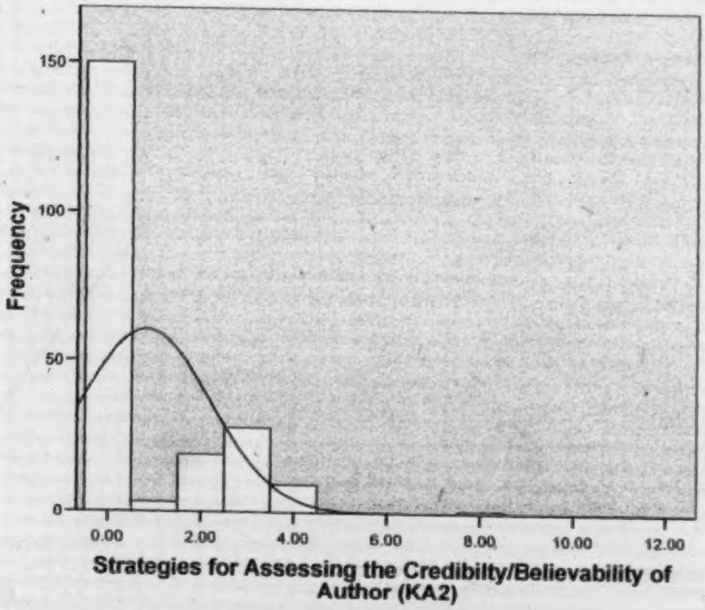

Figure 9. A histogram of participants' strategies for assessing the credibility or believability of the author of a Web resource (baseline sample) with normal curve.

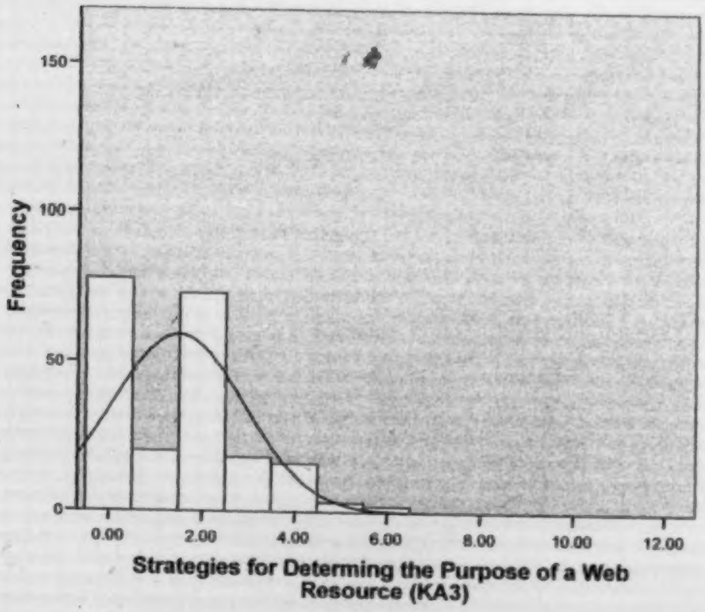

Figure 10. A histogram of participants' strategies for determining the purpose of a Web resource (baseline sample) with normal curve.

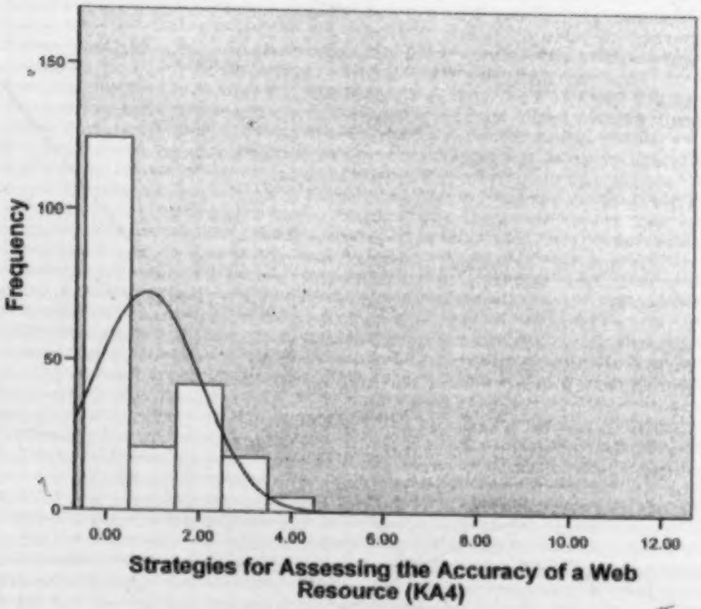

Figure 11. A histogram of participants' strategies for assessing the accuracy of a Web resource (baseline sample) with normal curve.

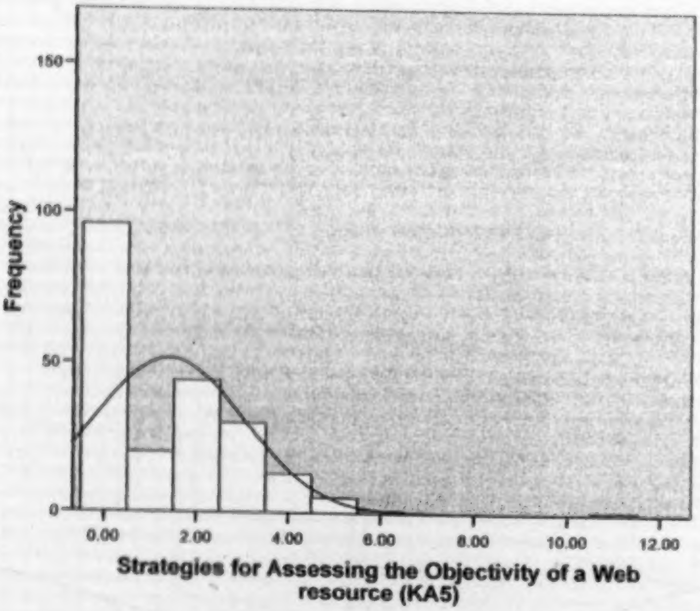

Figure 12. A histogram of participants' strategies for assessing the objectivity of a Web resource (baseline sample) with normal curve.

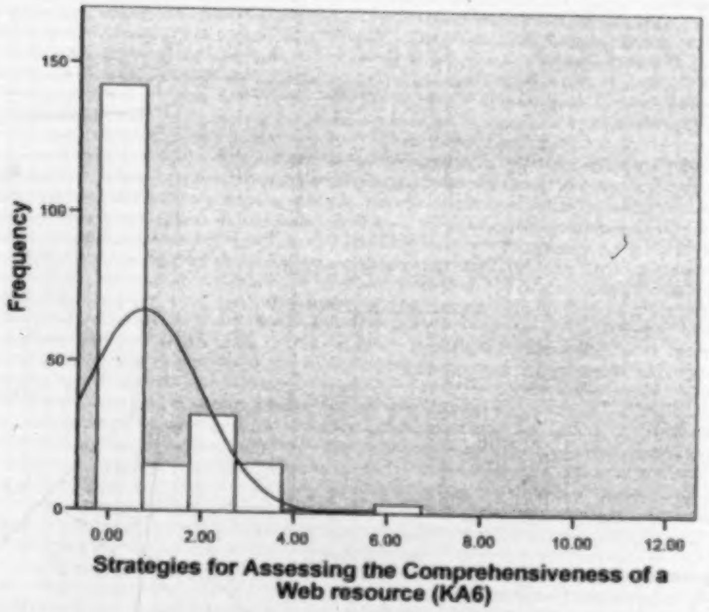

Figure 13. A histogram of participants' strategies for assessing the comprehensiveness of a Web resource (baseline sample) with normal curve.

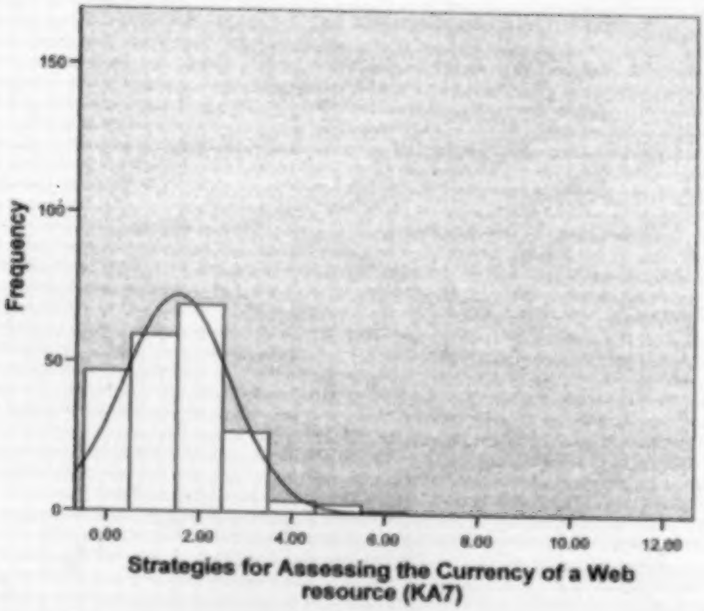

Figure 14. A histogram of participants' strategies for assessing the currency of a Web resource (baseline sample) with normal curve.

Motivation. Personal relevance (MO1) scores had a mean of 14.62 (out of 40) with a standard deviation of 6.85 (see Figure 15). Personal responsibility (MO2) had a mean score of 16.67 (out of a possible 35) with a standard deviation of 6.84 (see Figure 16). Motivation today (MO3) had a mean of 3.16 (out of 5) with a standard deviation of 1.02 (see Figure 17). Motivation in general (MO4) had a mean of 2.80 (out of 5) and a standard deviation of 1.00 (see Figure 18).

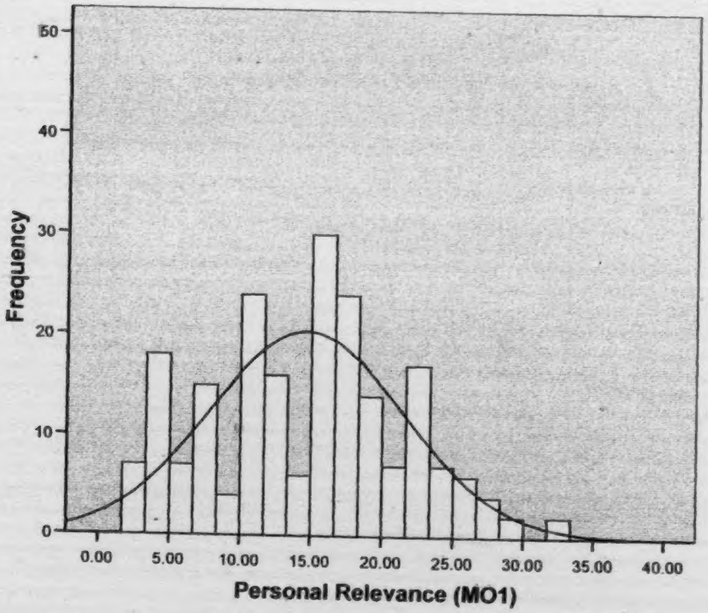

Figure 15. A histogram of participants' ratings of personal relevance (baseline sample) with normal curve.

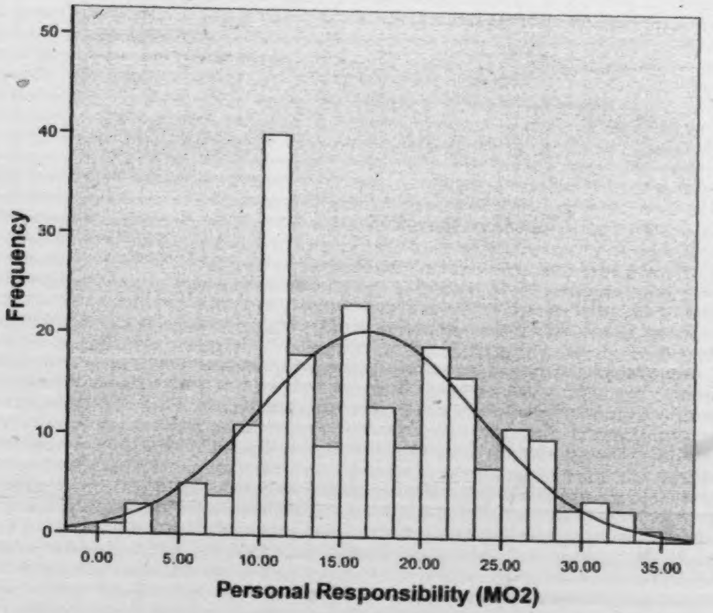

Figure 16. A histogram of participants' ratings of personal responsibility (baseline sample) with normal curve.

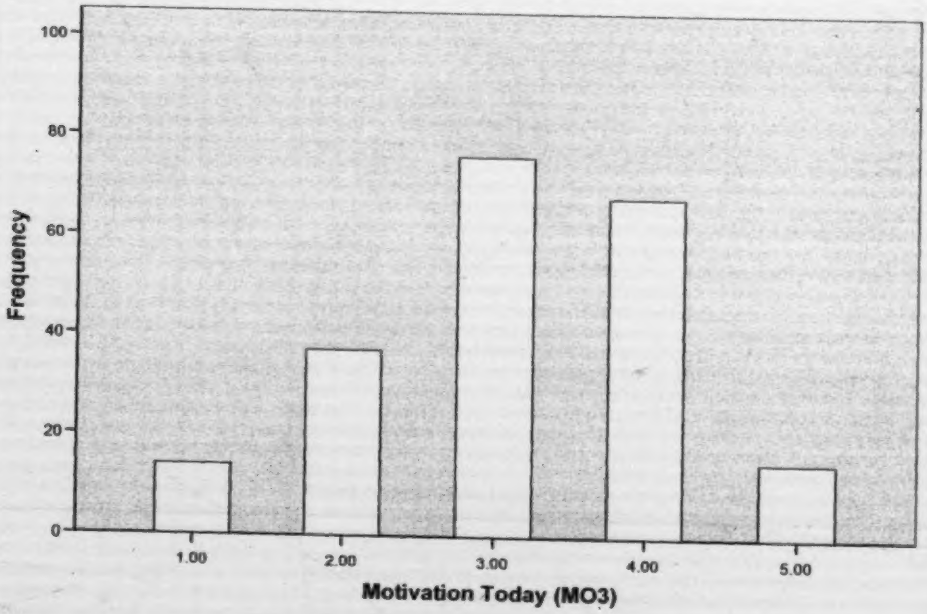

Figure 17. A bar graph of participants' ratings of their motivation on the day of testing (baseline sample) with normal curve.

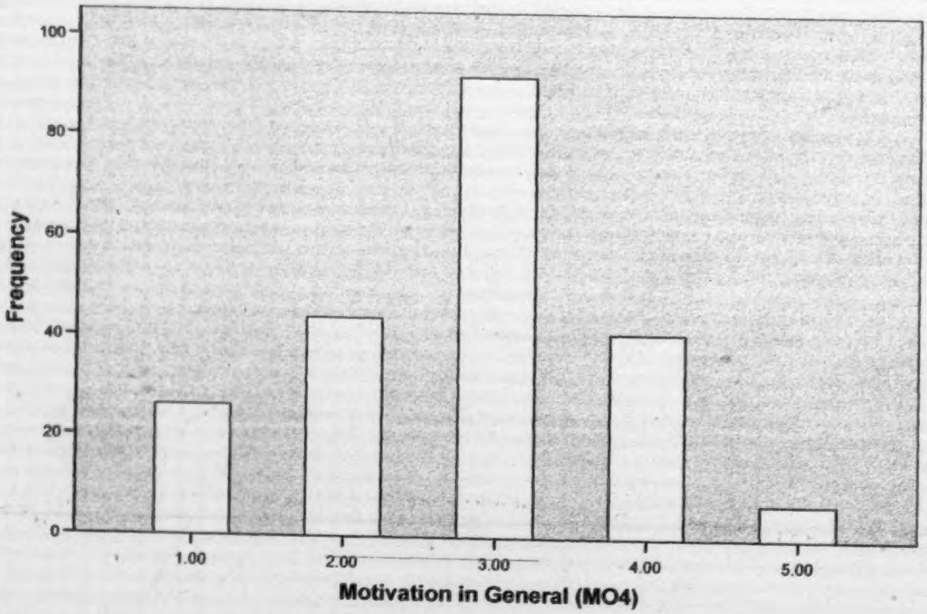

Figure 18. A bar graph of participants' ratings of their general motivation to ensure that information regarding alcohol abuse is accurate (baseline sample).

Critical Appraisal. Rating of the Web resource (CA1) was done using a 5-point likert-type scale. Mean response was 3.86 with a standard deviation of 0.79 (see Figure 19). Critical appraisal of the Author (CA2) had a mean of 7.58 (out of 13) with a standard deviation of 2.09 (see Figure 20). Critical appraisal of the purpose of the Web resource (CA3) had only 3 possible values (1, 2, or 3). The mean was 1.08 with a standard deviation of 0.28. Critical appraisal of the content of the resource (CA4) had a maximum score of 12 and the mean response was 7.30 with a standard deviation of 2.46 (see Figure 21). Critical appraisal of the form of the Web resource (CA5) had a mean of 5.43 and a standard deviation of 1.47 (out of a total of 9-see Figure 22).

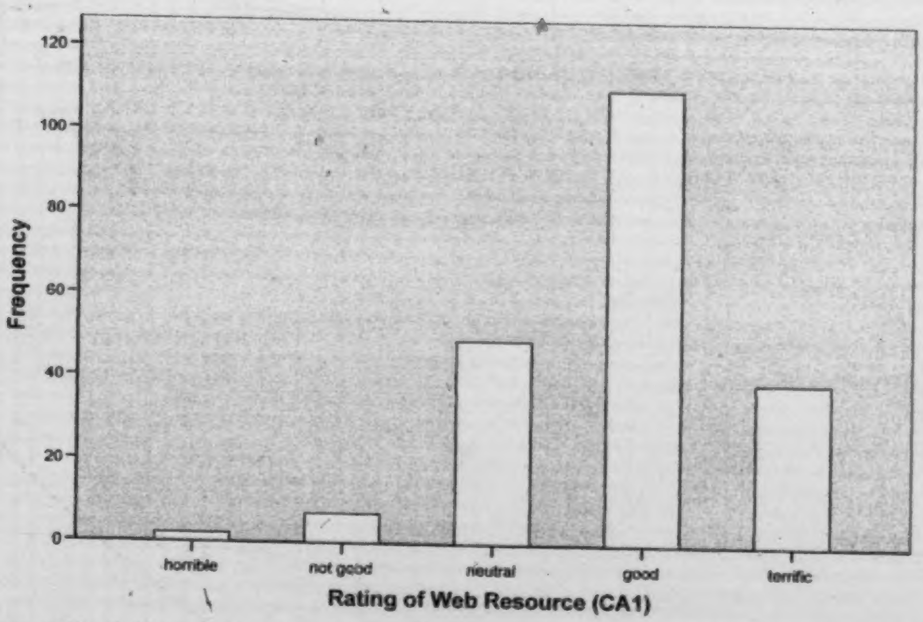

Figure 19. A bar graph of participants' ratings of the test Web resource (baseline sample).

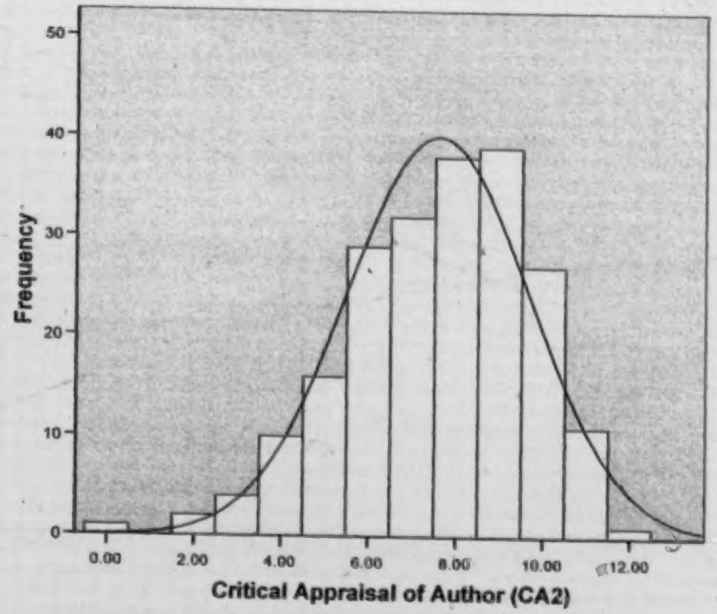

Figure 20. A histogram of participants' critical appraisal scores regarding the author of a Web resource (baseline sample) with normal curve.

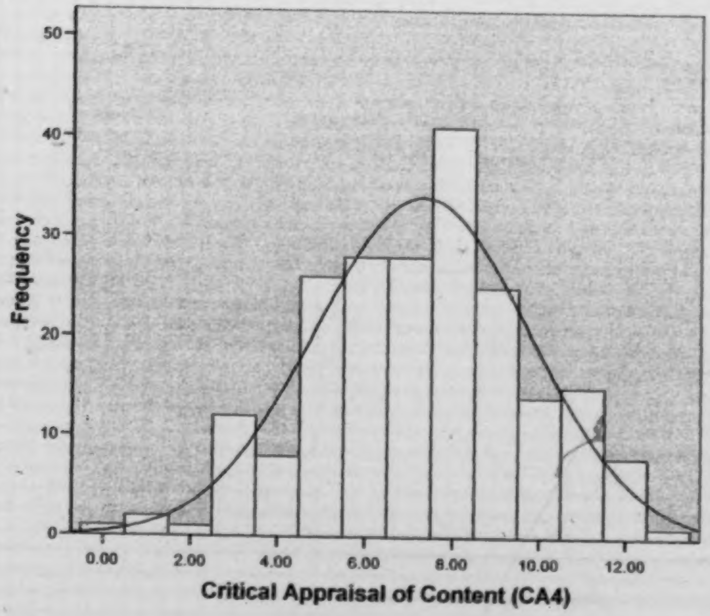

Figure 21. A histogram of participants' critical appraisal scores regarding the content of a Web resource (baseline sample) with normal curve.

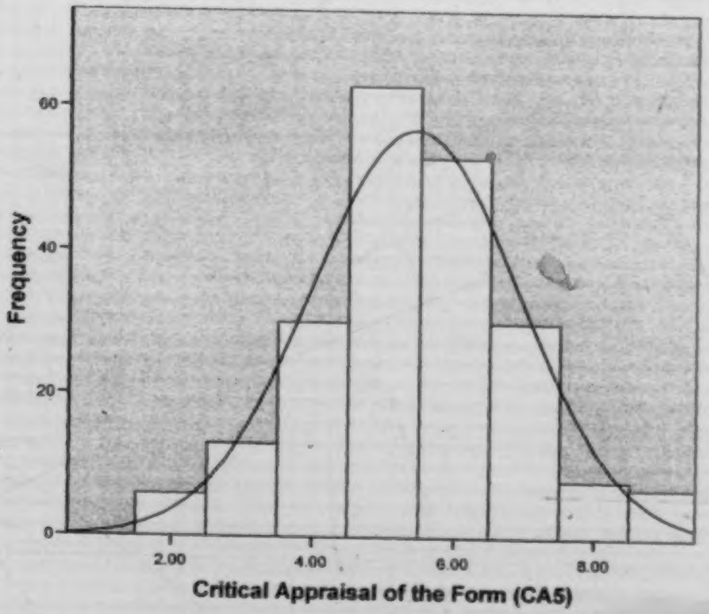

Figure 22. A histogram of participants' critical appraisal scores regarding the form of a Web resource (baseline sample) with normal curve.

Structural Analysis

A frequent problem in many SEM models is the occurrence of Heywood cases, or negative error variances. In examining the current model, there was a slight discrepancy as the variance for the disturbance term was negative (-.02). In order to deal with this problem, this variance term was fixed at zero and the model was rerun. This is a common practice for dealing with Heywood cases, and when the variance is small does not tend to cause too many problems (c.f., Zhu, Walter, Rosenbaum, Russell, & Rainer, 2006).

Some such as Kline, while acknowledging that this is a common solution caution against this approach as the bias in the interpretation tends to increase (Kline, 2006). Other research has examined the problem of Heywood cases and how to deal with them. Chen, Bollen, Paxton, Curran and Kirby (2001), for example have looked at some potential reasons for Heywood cases as they tend to be one of the most common types of problems associated with using SEM. Chen et al., conclude that we simply do not know enough yet to make claims as to why this problem occurs and what the consequences are. In addition, Chen et al., claim that researchers should not rely solely on negative error variance as an indicator of model misspecification. According to Chen et al., constraining Heywood cases to zero, while potentially biasing the results can actually lead to less bias than using an unconstrained model assuming that there are no problems with multicollinearity, no outliers, the negative error variance is not significant, and the negative error variance is small. The resulting, constrained model, was not significantly different from the original unconstrained model ($\chi^2(181, N = 210) = 240.63$ vs. $\chi^2(180, N = 210) = 240.52$ respectively). As a result, the constrained model is used from this point on in the analyses.

The first analysis performed was on the proposed full structural model (Figure 23) using the baseline sample. The resulting path analysis is shown in Figure 24.

The model is structurally identified, as there are 231 distinct sample moments and 49 sample parameters that need to be estimated. This results in 183 degrees of freedom. In addition, the model is recursive and appears to be empirically identified as their does not seem to be any multicollinearity nor do the start values seem to be problematic.

Before examining the fit of the model itself, it was necessary to examine the adequacy of the manifest variables in explaining the latent variables. As there are only two latent variables for domain knowledge, these were not sufficient to converge. To deal with this, the regression weighting for each of these were set to 1.00. In addition, because of the necessity of setting a scaling requisite, one of the parameters needed to be set at 1.00. For each factor, the loading parameter for the first variable is set to 1.00. As a result, there was no significance data reported for these parameters. From baseline sample 1, it is apparent that the majority of the manifest variables are adequate explanatory variables for their respective latent variables. The exceptions to this being: KA2, KA4, CA3, and CA5 (See Table 1).

The overall model fit was significant $\chi^2(181, N = 210) = 240.63, p < .05.$, however the relative chi-square ratio was 1.3, lower than the criteria of 2-5 stated by Kline (2005). The fit criteria associated with the model also seem to indicate that the model does fit the data moderately well with goodness-of-fit (GFI) = 0.90 when compared with a standard of 0.90 (Byrne, 2001; Hair et al., 1998). The comparative fit index (CFI) =0.84 also indicates a moderate fit when compared to a standard of 0.95 (Hu & Bentler, 1999). The root mean square error of approximation (RMSEA) = 0.04 is well below the established benchmark of 0.05 (Byrne, 2001; Diamantopoulos & Siguaw, 2000; Hair et al. 1998). This is further supported by the RMSEA 90% confidence interval = (.02, .05). A more extensive listing of fit measures can be found in Table 2. All of the measures indicate that the proposed model fits the data moderately well.

As shown in Figure 24, knowledge of critical appraisal is a significant predictor of critical appraisal with a standardized regression weight of 0.72 (C.R. = 2.13, $p < .05$). Domain knowledge (DK) with a standardized regression weight of 0.64 (C.R. = 1.85, $p > .05$) and motivation (MO) with a standardized regression weight of 0.52 (C.R. = 1.79, $p > .05$) while not significant do appear to be on the borderline. Computer/Web knowledge (CK) does not appear to be a significant predictor with a standardized regression weight of -0.01 (C.R. = -0.07, $p > .05$).

In addition to these main effects, the covariance (Table 3) between Computer/Web knowledge and Motivation was significant with a standardized estimate of 0.28 (C.R. = 2.32, $p < .05$). The covariance between knowledge of critical appraisal and motivation, while not significant, did approach significance having a standardized regression weight of -0.27 (C.R. = -1.91, $p > .05$). None of the other covariance measures appear to be significant.

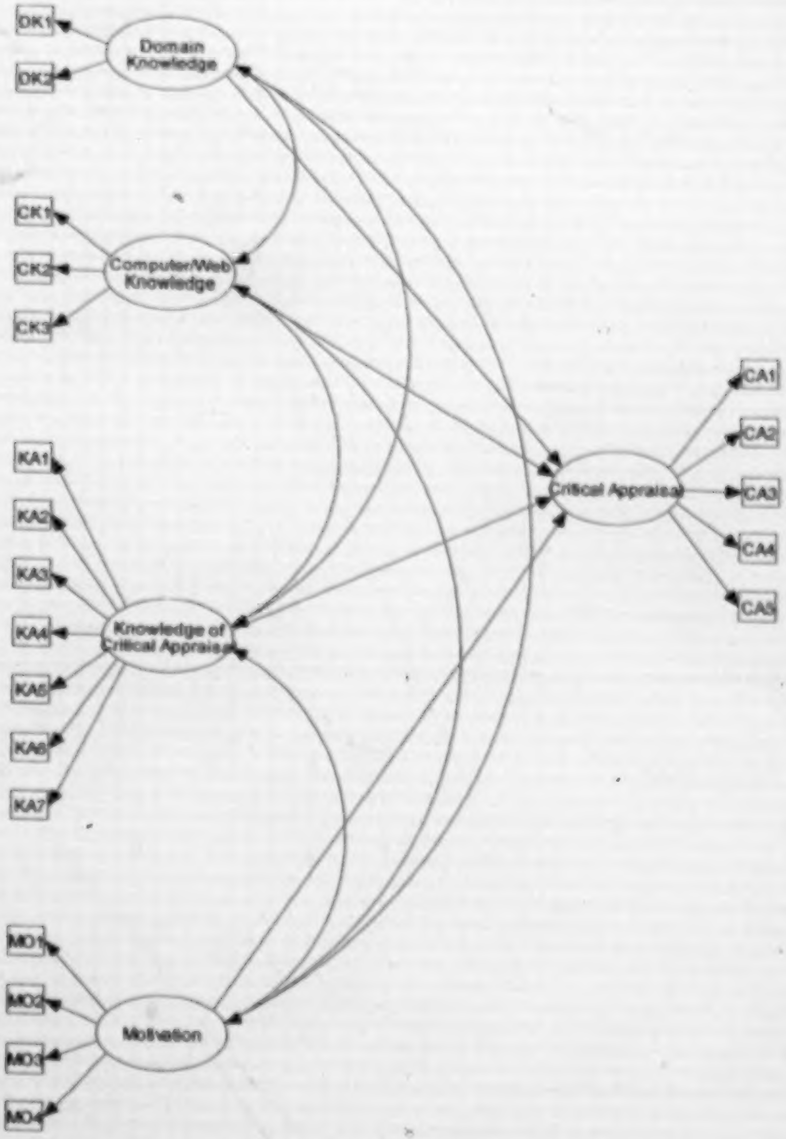
DK1
DK2
Domain Knowledge
CK1
CK2
CK3
Computer/Web Knowledge
KA1
KA2
KA3
KA4
KA5
KA6
KA7
Knowledge of Critical Appraisal
MO1
MO2
MO3
MO4
Motivation
Critical Appraisal
CA1
CA2
CA3
CA4
CA5

Figure 23. A proposed structural model including latent variables.

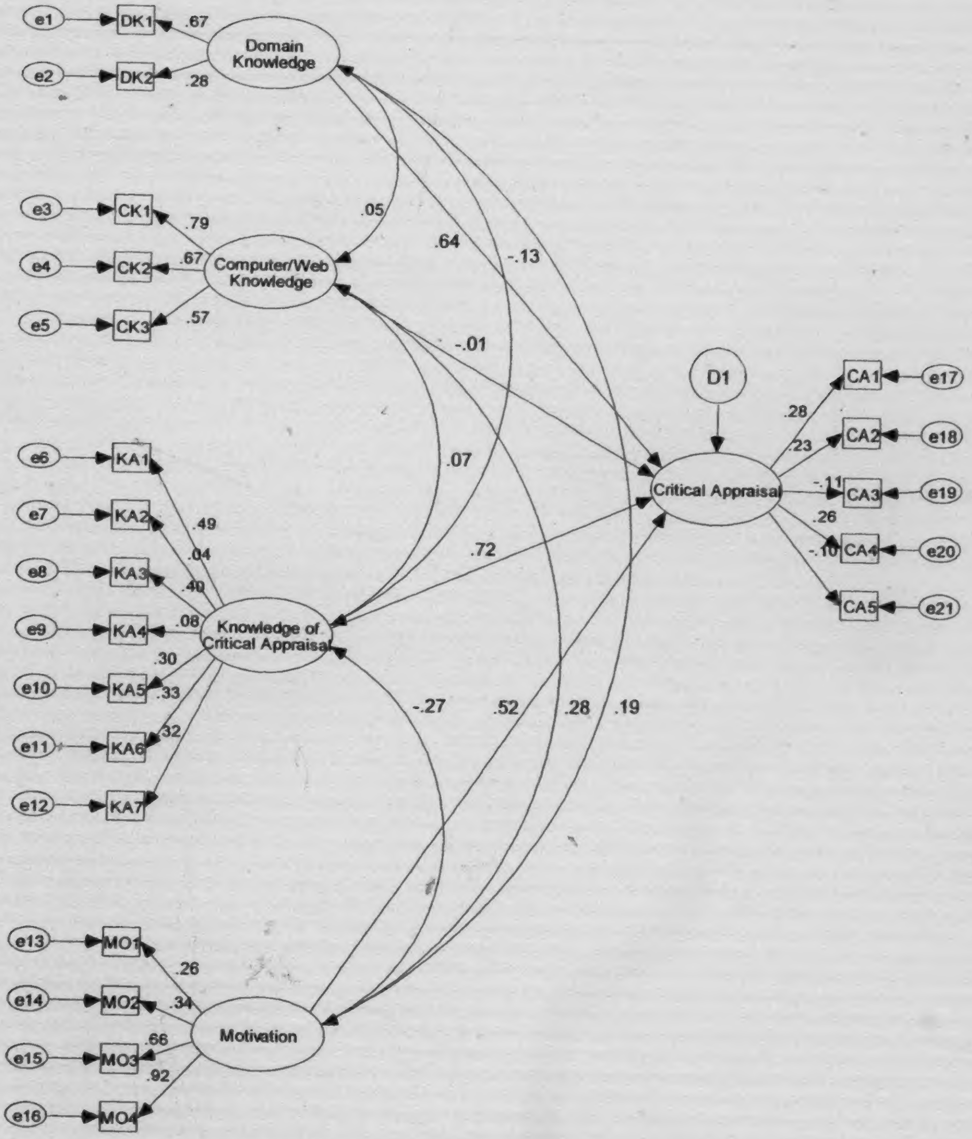
e1
DK1
.67
Domain Knowledge
e2
DK2
.28
e3
CK1
.79
e4
CK2
.67
Computer/Web Knowledge
e5
CK3
.57
.05
.64
-.13
-.01
D1
Critical Appraisal
CA1
e17
.28
CA2
e18
.23
CA3
e19
-.11
.26
CA4
e20
-.10
CA5
e21
.07
e6
KA1
e7
KA2
e8
KA3
e9
KA4
e10
KA5
e11
KA6
e12
KA7
.49
.04
.40
.08
.30
.33
.32
Knowledge of Critical Appraisal
.72
-.27
.52
.28
.19
e13
MO1
e14
MO2
e15
MO3
e16
MO4
.26
.34
.66
.92
Motivation

Figure 24. A proposed structural model including latent variables and standardized regression weights.

Table 1

Standardized Regression Weights of Parameter Estimates for Latent Variables (Full Baseline Model)

Parameter	Standardized Regression Weight	Critical Ratio	Significance
DK1	0.6745	-	-
DK2	0.2790	-	-
CK1	0.7935	-	-
CK2	0.6680	6.4806	0.0000
CK3	0.5727	6.1777	0.0000
KA1	0.2347	-	-
KA2	0.0420	0.4324	0.6654
KA3	0.3967	2.9596	0.0031
KA4	0.0835	0.8482	0.3963
KA5	0.2991	2.5381	0.0111
KA6	0.3320	2.7060	0.0068
KA7	0.3185	2.6405	0.0083
MO1	0.2590	-	-
MO2	0.3448	2.9217	0.0035
MO3	0.6593	3.4529	0.0006
MO4	0.9232	3.3295	0.0009
CA1	0.2755	-	-
CA2	0.2347	1.9600	0.0500
CA3	-0.1150	-1.1581	0.2468
CA4	0.2624	2.0860	0.0370
CA5	-0.1022	-1.0453	0.2959

Table 2

Summary of Fit Indices for Full Baseline Model

Fit Index	Baseline Structural Model
χ^2	240.6277
Root Mean Square Residual (RMR)	0.4033
Goodness-of-fit Index (GFI)	0.9029
Adjusted Goodness-of-fit Index (AGFI)	0.8761
Parsimony Adjusted Goodness-of-fit Index (PGFI)	0.7075
Normed Fit Index (NFI)	0.5883
Relative Fit Index (RFI)	0.5223
Incremental Fit Index (IFI)	0.8522
Tucker-Lewis Index (TLI)	0.8152
Comparative Fit Index (CFI)	0.8407
Parsimony Ratio (PRATIO)	0.8619
Parsimony-adjusted Normed Fit Index (PNFI)	0.5070
Parsimony-adjusted Comparative Fit Index (PCFI)	0.7246
Noncentrality Parameter Estimate (NCP)	59.6277
NCP lower bound	23.3215
NCP upper bound	104.0245
Minimum Discrepancy Function (FMIN)	1.1513
Population Discrepancy (F0)	0.2853
F0 lower bound	0.1116
F0 upper bound	0.4977
Root Mean Square Error of Approximation (RMSEA)	0.0397
RMSEA lower bound	0.0248
RMSEA upper bound	0.0524
Akaike Information Criterion (AIC)	340.6277

Browne-Cudeck criterion (BCC)	352.3924
Bayes Information Criterion (BIC)	660.2092
Consistent Akaike Information Criterion (CAIC)	557.9830
Expected cross validation index (ECVI)	1.6298
ECVI lower bound	1.4561
ECVI upper bound	1.8422
Hoelter .05 index	186
Hoelter .01 index	199

Table 3

Covariance Measures between Latent Variables (Full Baseline Model)

Covariance	Standardized Estimate	Critical Ratio	Significance
Domain knowledge ↔ Computer/Web knowledge	0.05	0.3972	0.6913
Domain knowledge ↔ Knowledge of critical appraisal	-0.13	-0.8525	0.3939
Domain knowledge ↔ Motivation	0.19	1.5765	0.1149
Computer/Web knowledge ↔ Knowledge of critical appraisal	0.07	0.5707	0.5682
Computer/Web knowledge ↔ Motivation	0.28	2.3227	0.0202
Knowledge of critical appraisal ↔ Motivation	-0.27	-1.9060	0.0567

Re-specification of the Baseline Model

In an attempt to create a more parsimonious fit, as well as to attempt to increase the fit of the borderline parameters of the baseline model, Domain knowledge (DK) and Motivation (MO), I re-specified the baseline model. It is important to point out that at this point, the process moves from a purely confirmatory approach to an exploratory approach. In practical applications of SEM, the most frequently used method combines confirmatory and exploratory approaches in a model development approach (Zhu et al., 2006).

The first parameter I removed was the effect of Computer/Web knowledge (CK) on Critical Appraisal (CA) ($p > .05$), as it was non-significant. In addition to this, I also removed the covariance between DK and CK ($p > .05$) as well as the covariance between CK and Knowledge of critical appraisal (KA) ($p > .05$). Based on modification indices, I added a direct effect between the DK2 and MO1. This led to a net gain of 2 degrees of freedom. The resulting structural model can be seen in Figure 25.

Before testing the actual model, I again tested the regression weights of the manifest parameter estimates for each of the latent variables to ensure that they were measuring the intended parameters. As there are still only two latent variables for domain knowledge, these still were not sufficient to converge. The regression weighting for each of these were set to 1.0. In addition, because of the necessity of setting a scaling requisite, one of the parameters needs to be set at 1.0. For each factor, the loading parameter for the first variable is set to 1.0. As a result, there were no significance values reported for these parameters. From Table 4, it is apparent that the majority of the manifest variables are adequate explanatory variables for their respective latent variables. The exceptions to this being: KA2, KA4, CA3, and CA5.

The resulting re-specified model with corresponding regression weights can be seen in Figure 26. The overall model is still significant $\chi^2(183, N = 210) = 227.71$, $p < .05$, although the model is approaching non-significance as the re-specified model is non-significant at $p < .01$. The fit criteria associated with the model indicate a moderate fit, GFI = 0.91, comparative fit index (CFI) = 0.88, root mean square error of approximation (RMSEA) = 0.03, RMSEA 90% confidence interval = (0.02, 0.05). A more extensive listing of fit measures can be found in Table 5.

As shown in Figure 26, knowledge of critical appraisal is a significant predictor of critical appraisal with a standardized regression weight of 0.74 (C.R. = 2.12, $p < .05$). Motivation in the re-specified model does appear to be a significant predictor with a standardized regression weight of 0.54 (C.R. = 1.96, $p < .05$). Domain knowledge (DK) with a standardized regression weight of 0.63 (C.R. = 1.86, $p > .05$) still is not significant but does remain on the borderline.

In addition to these main effects, the covariance (Table 6) between Computer/Web knowledge and Motivation was significant with a standardized estimate of 0.29 (C.R. = 2.46, $p < .05$. The covariance between knowledge of critical appraisal now appears to be significant with a standardized estimate of -0.29 (C.R. = -2.06, $p < .05$). None of the other covariance measures appear to be significant, however since removal of the remaining covariance measures weakened the fit.

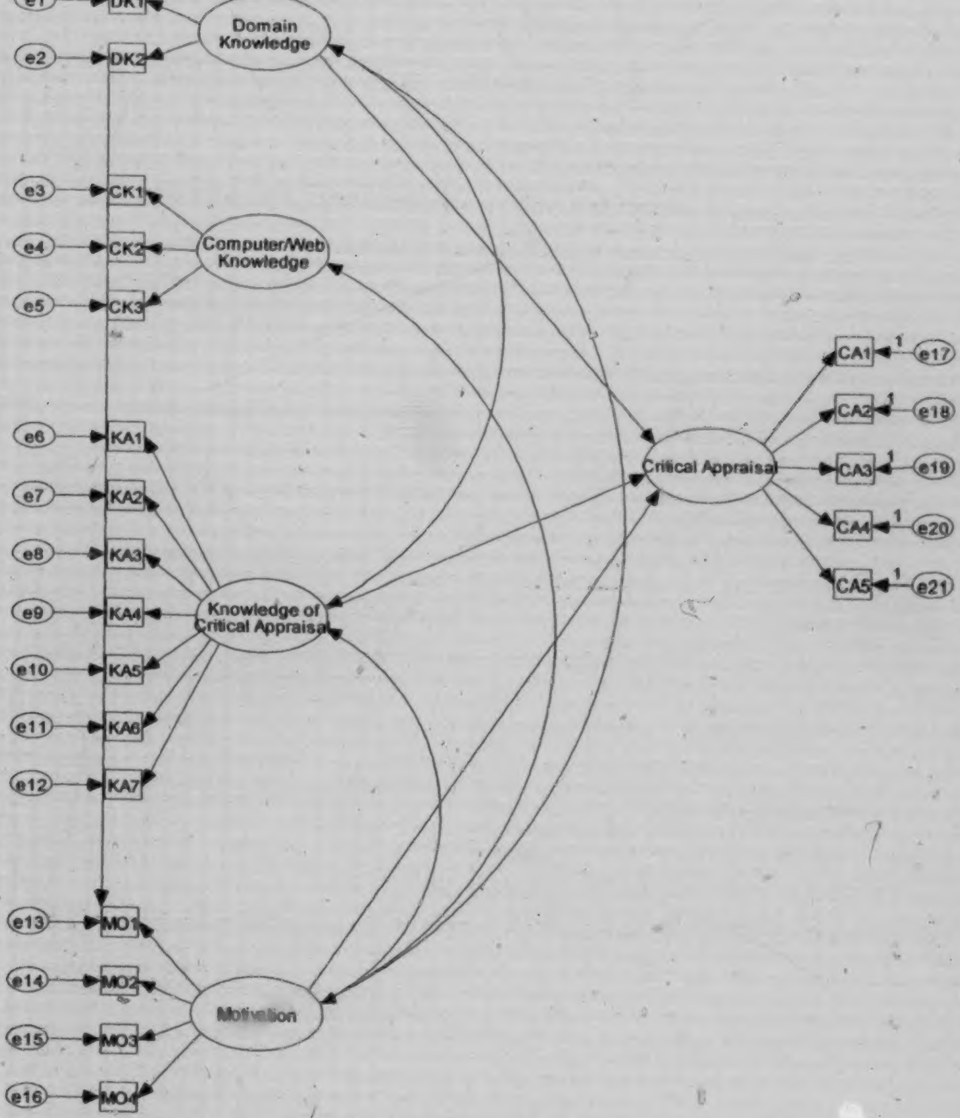
e1
e2
DK1
DK2
Domain Knowledge
e3
e4
e5
CK1
CK2
CK3
Computer/Web Knowledge
e6
e7
e8
e9
e10
e11
e12
KA1
KA2
KA3
KA4
KA5
KA6
KA7
Knowledge of Critical Appraisal
e13
e14
e15
e16
MO1
MO2
MO3
MO4
Motivation
Critical Appraisal
CA1
CA2
CA3
CA4
CA5
1
1
1
1
1
e17
e18
e19
e20
e21

Figure 25. Proposed re-specified structural model.

Table 4

Standardized Regression Weights of Parameter Estimates for Latent Variables (Re-specified Baseline Model)

Parameter	Standardized Regression Weight	Critical Ratio	Significance
DK1	0.6761	-	-
DK2	0.2796	-	-
CK1	0.7988	-	-
CK2	0.6663	6.4415	0.0000
CK3	0.5680	6.1258	0.0000
KA1	0.4842	-	-
KA2	0.0420	0.4314	0.6662
KA3	0.4025	2.9681	0.0030
KA4	0.0843	0.8544	0.3929
KA5	0.3025	2.5493	0.0108
KA6	0.3338	2.7064	0.0068
KA7	0.3143	2.6113	0.0090
MO1	0.2712	-	-
MO2	0.3453	3.1012	0.0019
MO3	0.6548	3.7450	0.0002
MO4	0.9299	3.5989	0.0003
CA1	0.2801	-	-
CA2	0.2294	1.9379	0.0526
CA3	-0.1150	-1.1562	0.2476
CA4	0.2596	2.0806	0.0375
CA5	-0.1056	-1.0738	0.2829
MO1 ← DK2	0.2436	3.7713	0.0002

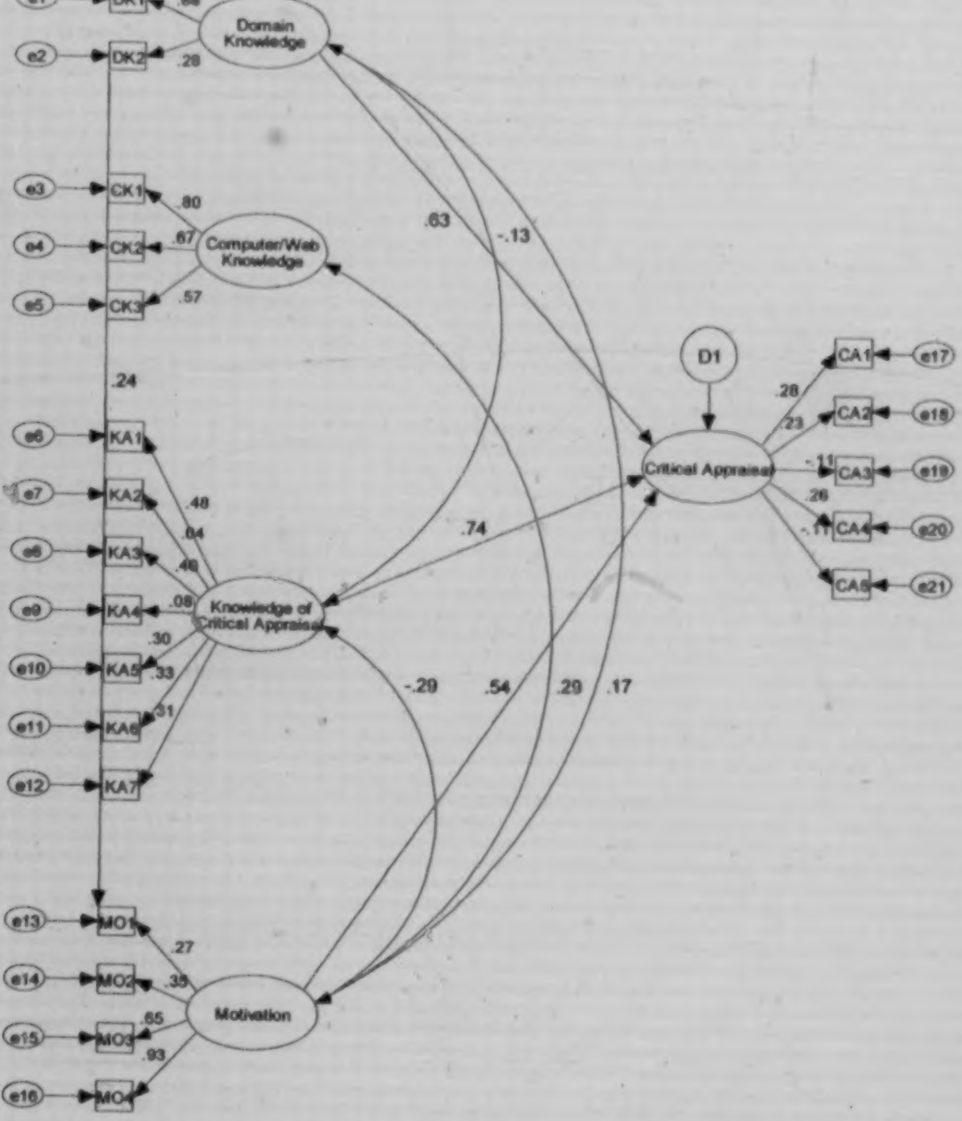
e1
e2
DK1
DK2
.68
.28
Domain Knowledge
e3
e4
e5
CK1
CK2
CK3
.80
.67
.57
Computer/Web Knowledge
.63
-.13
.24
e6
e7
e8
e9
e10
e11
e12
KA1
KA2
KA3
KA4
KA5
KA6
KA7
.48
.64
.46
.08
.30
.33
.31
Knowledge of Critical Appraisal
.74
D1
Critical Appraisal
CA1
CA2
CA3
CA4
CA5
e17
e18
e19
e20
e21
.28
.23
-.11
.26
-.11
-.29
.54
.29
.17
e13
e14
e15
e16
MO1
MO2
MO3
MO4
.27
.35
.65
.93
Motivation

Figure 26. Re-specified structural model with standardized regression weights.

Table 5

Summary of Fit Indices for Re-specified Baseline Model

Fit Index	Re-specified Baseline Structural Model
χ^2	227.7102
Root Mean Square Residual (RMR)	0.3413
Goodness-of-fit Index (GFI)	0.9071
Adjusted Goodness-of-fit Index (AGFI)	0.8827
Parsimony Adjusted Goodness-of-fit Index (PGFI)	0.7186
Normed Fit Index (NFI)	0.6104
Relative Fit Index (RFI)	0.5529
Incremental Fit Index (IFI)	0.8886
Tucker-Lewis Index (TLI)	0.8630
Comparative Fit Index (CFI)	0.8806
Parsimony Ratio (PRATIO)	0.8714
Parsimony-adjusted Normed Fit Index (PNFI)	0.5319
Parsimony-adjusted Comparative Fit Index (PCFI)	0.7674
Noncentrality Parameter Estimate (NCP)	44.7102
NCP lower bound	10.3573
NCP upper bound	87.2313
Minimum Discrepancy Function (FMIN)	1.0895
Population Discrepancy (F0)	0.2139
F0 lower bound	0.0496
F0 upper bound	0.4174
Root Mean Square Error of Approximation (RMSEA)	0.0342
RMSEA lower bound	0.0165
RMSEA upper bound	0.0478

Akaike Information Criterion (AIC)	323.7102
Browne-Cudeck criterion (BCC)	335.0043
Bayes Information Criterion (BIC)	630.5084
Consistent Akaike Information Criterion (CAIC)	532.3713
Expected cross validation index (ECVI)	1.5489
ECVI lower bound	1.3845
ECVI upper bound	1.7523
Hoelter .05 index	198
Hoelter .01 index	212

Table 6

Covariance Measures between Latent Variables (Re-specified Baseline Model)

Covariance	Standardized Estimate	Critical Ratio	Significance
Domain knowledge ↔ Knowledge of critical appraisal	-0.13	-0.8587	0.3905
Domain knowledge ↔ Motivation	0.17	1.5288	0.1263
Computer/Web knowledge ↔ Motivation	0.29	2.4636	0.0138
Knowledge of critical appraisal ↔ Motivation	-0.29	-2.0598	0.0394

Analysis of the Impact of Covariance Measures on the Structural Model

One of my primary goals in conducting this research was to develop and empirically valid model of critical appraisal of health-related Web resources. One component of this, and one of the motivations for using SEM was to examine the effects of covariance on the proposed model. While it is important to understand how both ability and motivation affect critical appraisal, a reductionist perspective is not the most effective technique. The original model was set up with the assumption that there were covariances between each of the contributing factors.

While parsimony is desirable, factors important for critical appraisal do not operate in a vacuum. In order to test this assumption, I compared the baseline model as constructed to this point with the same model, only having all of the covariance paths removed from the analysis. The resulting model can be seen in Figure 25.

From Table 7, it is apparent that the majority of the manifest variables are adequate explanatory variables for their respective latent variables. The exceptions to this being: KA2, KA4, CA3, and CA5. The resulting chi-square from the no-covariance baseline model is significant $\chi^2(187, N = 210) = 262.35, p < .05$. When compared with a the chi-square from the baseline model ($\chi^2(181, N = 210) = 240.63$, $p < .05$), the no-covariance baseline model is significantly less well fitting model than the baseline model.

The relative chi-square ratio was 1.4, which is still lower than the criteria of 2-5 recommended by Kline (2005). The fit criteria associated with the no-covariance baseline model, do seem to indicate that the model still fits the data moderately well with goodness-of-fit (GFI) = 0.90 when compared with a standard of 0.90 (Byrne, 2001; Hair et al., 1998). The comparative fit index (CFI) =0.71 also indicates a weaker fit when compared to a standard of 0.95 (Hu & Bentler, 1999). The root mean square error of approximation (RMSEA) = 0.04 is below the established benchmark of 0.05 (Byrne, 2001; Diamantopoulos & Siguaw, 2000; Hair et al. 1998). This is further illustrated by the RMSEA 90% confidence interval = (.03, .06). A more extensive listing of fit measures can be found in Table 2. Many of the measures indicate that the no-covariance baseline model fits the data relatively well.

As shown in Figure 25, knowledge of critical appraisal is a significant predictor of critical appraisal with a standardized regression weight of 0.63 (C.R. = 2.13, $p < .05$). Domain knowledge (DK) with a standardized regression weight of 0.54 (C.R. = 1.85, $p > .05$) and motivation (MO) with a standardized regression weight of 0.56 (C.R. = 1.79, $p > .05$) while not significant do appear to be on the borderline. Computer/Web knowledge (CK) still does not appear to be a significant predictor with a standardized regression weight of 0.06 (C.R. = -0.07, $p > .05$).

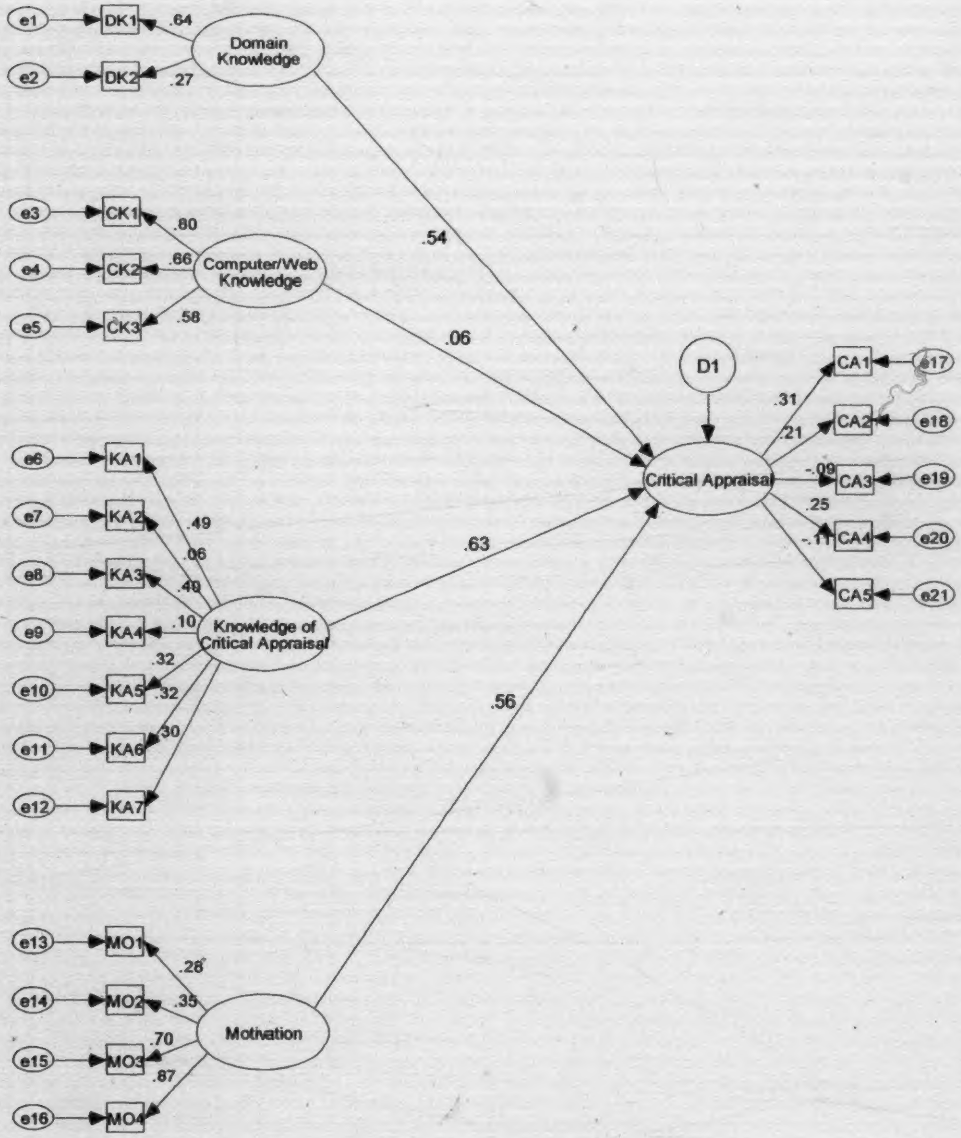
e1
DK1
.64
Domain Knowledge
e2
DK2
.27
e3
CK1
.80
e4
CK2
.66
Computer/Web Knowledge
e5
CK3
.58
.54
.06
D1
CA1
e17
.31
CA2
e18
.21
Critical Appraisal
-.09
CA3
e19
.25
-.11
CA4
e20
CA5
e21
e6
KA1
e7
KA2
.49
.06
e8
KA3
.49
.10
e9
KA4
Knowledge of Critical Appraisal
.63
.32
e10
KA5
.32
.30
e11
KA6
e12
KA7
.56
e13
MO1
.28
e14
MO2
.35
Motivation
.70
e15
MO3
.87
e16
MO4

Figure 27. Baseline model with covariance measures removed.

Table 7

Standardized Regression Weights of Parameter Estimates for Latent Variables (No-Covariance baseline Model)

Parameter	Standardized Regression Weight	Critical Ratio	Significance
DK1	0.642	-	-
DK2	0.2690	-	-
CK1	0.802	-	-
CK2	0.657	6.085	0.0000
CK3	0.576	5.924	0.0000
KA1	0.488	-	-
KA2	0.057	0.563	0.573
KA3	0.402	2.769	0.006
KA4	0.104	1.016	0.310
KA5	0.320	2.504	0.012
KA6	0.319	2.500	0.012
KA7	0.304	2.430	0.015
MO1	0.277	-	-
MO2	0.348	2.955	0.003
MO3	0.700	3.555	0.000
MO4	0.870	3.410	0.001
CA1	0.311	-	-
CA2	0.207	1.959	0.050
CA3	-0.093	-1.010	0.313
CA4	0.250	2.225	0.026
CA5	-0.122	-1.188	0.235

Table 8

Summary of Fit Indices for No-Covariance Baseline Model

Fit Index	Baseline Structural Model
χ^2	262.352
Root Mean Square Residual (RMR)	0.421
Goodness-of-fit Index (GFI)	0.895
Adjusted Goodness-of-fit Index (AGFI)	0.871
Parsimony Adjusted Goodness-of-fit Index (PGFI)	0.725
Normed Fit Index (NFI)	0.551
Relative Fit Index (RFI)	0.496
Incremental Fit Index (IFI)	0.810
Tucker-Lewis Index (TLI)	0.774
Comparative Fit Index (CFI)	0.799
Parsimony Ratio (PRATIO)	0.890
Parsimony-adjusted Normed Fit Index (PNFI)	0.491
Parsimony-adjusted Comparative Fit Index (PCFI)	0.711
Noncentrality Parameter Estimate (NCP)	75.352
NCP lower bound	36.579
NCP upper bound	122.160
Minimum Discrepancy Function (FMIN)	1.255
Population Discrepancy (F0)	0.361
F0 lower bound	0.175
F0 upper bound	0.584
Root Mean Square Error of Approximation (RMSEA)	0.044
RMSEA lower bound	0.031
RMSEA upper bound	0.056
Akaike Information Criterion (AIC)	350.352

Browne-Cudeck criterion (BCC)	360.705
Bayes Information Criterion (BIC)	631.584
Consistent Akaike Information Criterion (CAIC)	541.625
Expected cross validation index (ECVI)	1.676
ECVI lower bound	1.491
ECVI upper bound	1.900
Hoelter .05 index	176
Hoelter .01 index	188

Analysis of the Impact of Motivation as a Moderating Variable

An integral component of this analysis is the examination of motivation. Of the contributing factors that have been hypothesized, motivation is one that perhaps most often overlooked. One of the reasons for this may be due to the difficulty in operationalizing motivation. As described in this study, I defined motivation as consisting of elements of personal relevance, personal responsibility, current attitudes, and a general sense of motivation. In order to examine the effects of motivation on the other variables, I wanted to test the assumption that motivation acts as a moderating variable and that it influences the proposed ability factors. The resulting model can be seen in Figure 28.

From Table 9, it is apparent that the majority of the manifest variables are adequate explanatory variables for their respective latent variables. The exceptions to this being: KA2, KA4, CA3, and CA5. The resulting chi-square from the no-covariance baseline model is significant $\chi^2(184, N = 210) = 261.10, p < .05$. When compared with a the chi-square from the baseline model ($\chi^2(181, N = 210) = 240.63$, $p < .05$), the no-covariance baseline model is significantly less well fitting model than the baseline model.

The relative chi-square ratio was 1.4, which is still lower than the criteria of 2-5 recommended by Kline (2005). The fit criteria associated with the no-covariance baseline model, indicate that the model still fits the data moderately well with goodness-of-fit (GFI) = 0.90 when compared with a standard of 0.90 (Byrne, 2001; Hair et al., 1998). The comparative fit index (CFI) =0.79 also indicates a weaker fit when compared to a standard of 0.95 (Hu & Bentler, 1999). The root mean square error of approximation (RMSEA) = 0.05 is the same as the established benchmark of 0.05 (Byrne, 2001; Diamantopoulos & Siguaw, 2000; Hair et al. 1998). The RMSEA 90% confidence interval is (.03, .06). A more extensive listing of fit measures can be found in table 10. Many of the measures indicate that this model that has all of the motivation covariance measures removed fits the data relatively well, but significantly less well than the re-specified model.

As shown in Figure 28, knowledge of critical appraisal with a standardized regression weight of 0.68 (C.R. = 2.16, $p < .05$) and motivation (MO) with a standardized regression weight of 0.53 (C.R. = 2.14, $p < .05$) are significant predictors of critical appraisal. Domain knowledge (DK) with a standardized regression weight of 0.61 (C.R. = 1.87, $p > .05$) while not significant did appear to be on the borderline. Computer/Web knowledge (CK) still does not appear to be a significant predictor with a standardized regression weight of 0.02 (C.R. = 0.04, $p > .05$).

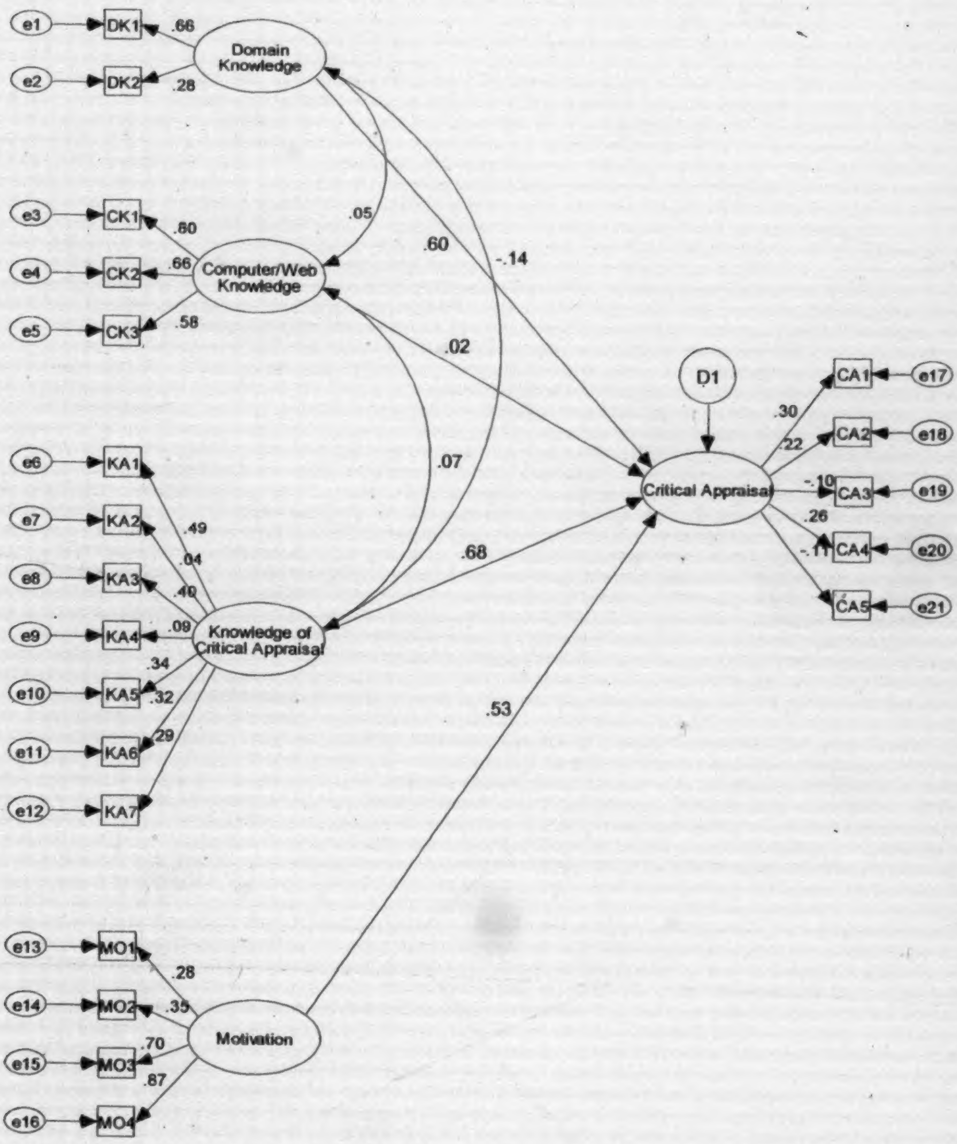
e1
DK1
.66
Domain Knowledge
e2
DK2
.28
e3
CK1
.60
.05
.60
-.14
e4
CK2
.66
Computer/Web Knowledge
e5
CK3
.58
.02
D1
CA1
e17
.30
CA2
e18
.22
e6
KA1
.07
Critical Appraisal
-.10
CA3
e19
.26
e7
KA2
.49
CA4
e20
-.11
.04
e8
KA3
.68
CA5
e21
.40
e9
KA4
.09
Knowledge of Critical Appraisal
.34
e10
KA5
.32
.53
e11
KA6
.29
e12
KA7
e13
MO1
.28
e14
MO2
.35
Motivation
.70
e15
MO3
.87
e16
MO4

Figure 28. Baseline model with motivation covariance measures removed.

Table 9

Standardized Regression Weights of Parameter Estimates for Latent Variables (Motivation Covariance Measures Removed)

Parameter	Standardized Regression Weight	Critical Ratio	Significance
DK1	0.664	-	-
DK2	0.276	-	-
CK1	0.795	-	-
CK2	0.659	6.143	0.0000
CK3	0.581	5.992	0.0000
KA1	0.490	-	-
KA2	0.040	0.405	0.685
KA3	0.400	2.797	0.005
KA4	0.094	0.928	0.353
KA5	0.335	2.591	0.010
KA6	0.335	2.514	0.012
KA7	0.294	2.404	0.016
MO1	0.277	-	-
MO2	0.348	2.945	0.003
MO3	0.700	3.553	0.000
MO4	0.870	3.403	0.001
CA1	0.301	-	-
CA2	0.216	1.960	0.050
CA3	-0.098	-1.036	0.300
CA4	0.259	2.195	0.028
CA5	-0.115	-1.187	0.235

Table 10

Summary of Fit Indices for Baseline Model (Motivation Covariance Measures Removed)

Fit Index	Baseline Structural Model
χ^2	261.097
Root Mean Square Residual (RMR)	0.421
Goodness-of-fit Index (GFI)	0.896
Adjusted Goodness-of-fit Index (AGFI)	0.869
Parsimony Adjusted Goodness-of-fit Index (PGFI)	0.713
Normed Fit Index (NFI)	0.553
Relative Fit Index (RFI)	0.490
Incremental Fit Index (IFI)	0.807
Tucker-Lewis Index (TLI)	0.765
Comparative Fit Index (CFI)	0.794
Parsimony Ratio (PRATIO)	0.876
Parsimony-adjusted Normed Fit Index (PNFI)	0.485
Parsimony-adjusted Comparative Fit Index (PCFI)	0.696
Noncentrality Parameter Estimate (NCP)	77.097
NCP lower bound	38.287
NCP upper bound	123.931
Minimum Discrepancy Function (FMIN)	1.249
Population Discrepancy (F0)	0.369
F0 lower bound	0.183
F0 upper bound	0.593
Root Mean Square Error of Approximation (RMSEA)	0.045
RMSEA lower bound	0.032
RMSEA upper bound	0.057

Akaike Information Criterion (AIC)	355.097
Browne-Cudeck criterion (BCC)	366.156
Bayes Information Criterion (BIC)	655.504
Consistent Akaike Information Criterion (CAIC)	559.411
Expected cross validation index (ECVI)	1.699
ECVI lower bound	1.513
ECVI upper bound	1.923
Hoelter .05 index	174
Hoelter .01 index	186

Discussion

The descriptive analysis revealed a number of important factors; first, when looking at the domain expertise scores, not surprisingly, the scores tend to cluster about the midpoint. Out of 15 possible, participants were able to answer correctly just over half of the items. This is not surprising considering that this was a non-specialized sample. The resulting normal distribution could describe quite a few different phenomena in the general population. In addition, participants' scores on the test of domain knowledge seem to match very well with their own reported domain knowledge.

Interestingly, when computer/Web knowledge and skill are measured, there is a definite pattern to the results. All of the measures (how much they like computers/Web, how frequently the use computers/Web, and how much skill they report having with computers/Web) are on the upper end of the scale. Two of the measures (how much they like computer/Web and how much skill they report having) are skewed to the left and appear to be truncated at the top end (see Figure 5 & Figure 7).

What is, perhaps, the most surprising are the scores for knowledge of critical appraisal. All of the mean responses are incredibly low (all less that 3). This is especially concerning when one considers that there is theoretically no upper end to the scores that people could get. In fact, not one person out of all 210 participants actually scored higher than 11 and the modal response in every one of the indicators was either 0 or 2.

Motivation scores were in the low to mid range. Personal relevance was slightly lower that personal responsibility but both of the responses were near the mid-range. Both motivation today and motivation in general also showed similar patterns with motivation today being slightly higher than general motivation.

Critical appraisal measures were a little more varied. The rating of the Web resource was positive. Critical appraisal of the author was slightly higher than the mid-range. Critical appraisal of the purpose of the Web resource was slightly below the mid-range and critical appraisal of the content was slightly above the mid-range. Critical appraisal of the form was also fairly high.

The overall model that was tested, using alcohol abuse as the domain, was not significant according to the chi-square goodness of fit measure. While this might cause some concern regarding the resulting utility of the proposed model to explain the resulting variability, there are some potential problems associated with relying solely on the chi-square goodness of fit measure (Bentler & Bonet, 1980; Byrne, 2001). Specifically, the concern is that the chi-square is overly sensitive because of the large sample size.

In addition to the chi-square, it is common practice to also look at a number of other fit measures in order to determine whether the model fit is within acceptable parameters. While this does not lend itself as neatly to the traditional logic of hypothesis testing, it does allow for an understanding of the probable utility of the model (Byrne, 2001).

When the additional fit measures are examined, it is clear that the proposed model does show some promise as a potential explanation for the data. For example, the obtained GFI is well within accepted parameters. GFI values range from 0 to 1 with values that are closer to 1 indicating a good fit (Byrne, 2001; Hair et al., 1998).

In addition to the GFI, Bentler (1990) claims that the CFI is also an important measure to examine. The obtained value is below the value of 0.95 which is accepted as an indication of a good fitting model (Bentler, 1990; Hu & Bentler, 1999). The closer that a value is to this cutoff of 0.95 indicates how well the model fits the data. The obtained value, while not perfect, can be considered an indication of a moderate fit.

Not only is it important to consider the GFI and the CFI when determining fit of the proposed model. The RMSEA is also recommended as an important value that should be considered when evaluating a potential structural model (Byrne, 2001; Diamatopoulos & Siguaw, 2000; Hair et al., 1998). The obtained value is another indication of the potential utility of this model in describing the data. The resulting 90% confidence interval of is further support for this particular measure.

The conclusion that can be drawn from examination of this structural model is that the proposed model is moderately well fitting. There is potential for this model to explain the data and perhaps to allow for an examination of the factors that are important for examining critical appraisal. The factors that are important for predicting critical appraisal are domain knowledge, knowledge of critical appraisal and motivation.

The fact that domain knowledge is an important predictor of critical appraisal is, perhaps, not surprising. Having a high level of domain knowledge can be important for a number of reasons. For example, domain knowledge does allow one to understand the key issues related to a particular topic. The effect of this knowledge is an increased capacity to chunk information, and more efficiently process domain specific information. As a result, a user may be more likely to notice key features that allow the user to make informed decisions. As Fogg et al. (2002) have found, domain experts are more likely to make decisions regarding the trustworthiness of a Web resource based on more than how it looks.

Perhaps another reason why domain knowledge is important in determining whether someone will critically appraise a Web resource is the fact that domain knowledge may be related to motivation. As seen from the re-specified model (see Table 6), there does seem to be a significant covariance between domain knowledge and motivation. In addition to this, motivation is another significant predictor of critical appraisal. Whether someone takes the time and effort to critically appraise a Web resource does seem to be tied to how motivated that person is to critically appraise.

Motivation can come from many different sources. The factors that were considered in this particular study were personal relevance, personal responsibility, general motivation and motivation to perform the task. Each of the different factors that were measured in this study contributed to the overall measure of motivation. If a person is motivated enough, there is evidence to suggest that this person will be more likely to expend the energy that it takes to critically appraise a Web resource.

As research by Simon (1957) has shown, effortful processing does not always take place. If it is assumed that there is a limited amount of cognitive resources available for a given task (Navon & Gopher, 1979), it may also be assumed that procedures that reduce the need for these cognitive resources would be preferred over other more cognitively intensive procedures (Macrae, Milne, & Bodenhausen, 1994), unless there is good reason to expend these cognitive resources. As can be seen from this structural model, increased motivation is, in fact, related to the likelihood that one will critically appraise a Web resource.

Finally, in addition to domain knowledge and motivation, the effect of knowledge of critical appraisal is also a significant predictor of critical appraisal. Knowing what one should look for in order to make an informed decision regarding a Web resource is important because without this knowledge it would be difficult, if not impossible to make an informed decision. As was seen in this study, knowledge of critical appraisal is very low. A consequence of this is that it may actually function as a bounding condition and might be one of the reasons that people do not critically appraise on a regular basis.

Interestingly, knowledge of computers and the Web does not seem to significantly predict the likelihood for critical appraisal. In retrospect, this is, perhaps, not surprising. What became clear as the results were examined is that, for this sample, there appeared to be somewhat of a ceiling effect for knowledge of the Web. Most of the self-ratings of Web and computer knowledge were very high. So there really was no clear group of participants who were low in Web and computer knowledge. The effect of this is that it perhaps really does not allow for a clear picture of what the effect of Web and computer knowledge would be. Or alternatively, that it was not measured appropriately.

While Web and computer knowledge do not directly predict critical appraisal, this parameter does have a significant relationship with two of the other parameters, knowledge of critical appraisal and motivation. The reasons for these relationships are not immediately apparent; however, it is possible that knowing how to use the computer has a global effect on the motivation to perform a task and on the ability that one has to focus on factors of critical appraisal that may otherwise be otherwise overlooked.

As the baseline model did show some promise and it does have some exciting implications for explaining the critical appraisal of health-related Web resources, I believed that it was still be possible to continue to improve on this model. One of the benefits of structural equation modeling is that it allows for re-specification of structural models in order to improve the fit of a model (Zhu et al., 2006). In an attempt to increase the fit of the model and to improve upon the parsimony (without removing all of the measures of covariance) of the baseline model, I re-specified the baseline model.

The re-specified model was an improvement over the baseline model. This re-specified model was still significant according to the chi-square analysis; however, the chi-square is closer to non-significance than it was in the baseline model. The fit criteria that were used to evaluate the baseline model also indicate that the re-specified model is a better fitting model (compare Table 2 and Table 5). The GFI of the re-specified model is a slightly stronger indicator of goodness of fit than the GFI of the baseline model. Likewise, the CFI measure of the re-specified model is a slightly stronger indicator of goodness-of-fit than the CFI of the baseline model. By comparing the RMSEA values of the original structural model and the re-specified model as well as the 90% confidence interval, it is again possible to see that the re-specified model is a better fit. Working through each of the other fit parameters and finding similar patterns strengthens the argument that the re-specified model is an improvement over the original baseline model.

Part of the process involved in re-specifying the model was to determine which parameters, if any, were not contributing to the overall strength of the model. Additionally it is possible that some parameters could actually weaken the proposed model. So in order to ensure that the final model is the best fitting model possible, all of the contributing parameters and associated measures of covariance were examined to determine whether there were some that could be removed.

One of the considerations involved in removing or otherwise re-specifying a structural model, is that once one moves to this stage, one enters into an exploratory approach (Byrne, 2001). Correspondingly, the possibility of chance findings and spurious results does become a concern. It was with this cautionary note that that the re-specification was completed.

The first thing that became obvious when examining the baseline model is that knowledge of computers and the Web is not a significant predictor of critical appraisal. While this parameter was clearly not a direct predictor of critical appraisal, one of the strengths of structural equation modeling is the capability to examine measures of covariance in addition to the main effects. Examining the measures of covariance indicated that there were significant measures of covariance between knowledge of computers and the Web and both motivation and knowledge of critical appraisal. As such, it did not seem wise to exclude this parameter completely. Instead, the effect of the parameter on critical appraisal was simply removed while the significant covariance was allowed to continue to affect the corresponding parameters. The most significant contributing factors in determining whether critical appraisal will take place are, first, one's knowledge of critical appraisal, second one's domain knowledge and third the motivation that one feels to critically appraise.

As can be seen from Table 6, the covariance between computer and Web knowledge and motivation is significant. Computer and Web knowledge does not directly affect critical appraisal, but apparently it does have an effect on the motivation that one experiences to critically appraise. So perhaps having a high level of computer knowledge increases confidence levels and allows one to be more curious or to be more motivated to try and critically appraise information that one comes into contact with.

Another covariance that was significant was the relationship between knowledge of critical appraisal and motivation. One of the interesting things about this relationship is that it does appear to be a negative relationship indicating that as knowledge of critical appraisal increases, there is a corresponding decrease in the levels of reported motivation. Perhaps there is a relationship between how much one knows about the different factors that allow one to critically appraise, and how willing one is to engage in critical appraisal.

In addition to these two, significant, measures of covariance, there are also two other measures of covariance, that while not significant, did seem to increase the predictive power of the proposed model. The first of these was the relationship between domain knowledge and knowledge of critical appraisal. Once again, there appears to be a negative relationship. Increased domain knowledge is related to a decrease in the knowledge of critical appraisal. Because there is no way to test the directionality of this relationship, it is impossible to say which of these factors impacts the other, or whether there is some third factor or set of factors that impacts both of the separate parameters.

There is also a relationship, although non-significant between domain knowledge and motivation. This is a positive relationship and although non-significant still does appear to have an effect. The more domain knowledge, the more motivated people are, or alternatively the more motivation people have, the more domain knowledge they acquire.

These relationships are critical because they point to the fact that while there are significant direct predictors of critical appraisal, there are also indirect influences that can affect each other. As noted in the introduction, one of the reasons why this is critical research to perform is because some of the current models, or hypothesized relationships between individual patterns of Web searching, assume a very simplistic relationship (c.f., Lazonder, 2000).

To test whether it was useful to include the measures of covariance in the model, I created a no-covariance baseline model. The analysis of this model indicated that the original baseline model was a significantly better fitting model than was the no-covariance baseline.

Although the baseline model was better fitting than the no-covariance baseline model, the no-covariance baseline model still seemed to fit the data relatively well according to the fit indices. What this, perhaps, indicates is that while the baseline model fits the data better than the no-covariance baseline model, it is apparent that it might be useful to remove at least some of the measures of covariance in order to continue to improve fit.

In addition, the fact that the baseline model is a better fitting model supports the notion that it is important to consider not just the individual factors and how they influence the critical appraisal of health-related Web resources, but also the potential measures of covariance between them.

It is important to note that the no-covariance model did not impose any sequential constraints upon the contributing factors. The fact that this model is still moderately well fitting, and potentially useful, but not as well fitting as the models with the covariances, lends support to the contention that it is important not just to consider each factor. All factors need to be examined simultaneously and it is important to take into account that there are many different influences with respect to the critically appraisal of health-related Web resources. Another assumption that I tested in this research was that motivation plays a moderating role with the ability factors. What this means is that the influence of ability factors on the likelihood of critical appraisal is influenced by motivation. I tested this assumption by testing a version of the model in which I removed all covariance measures between motivation and other factors.

The resulting model significantly weaker than the original baseline model. This supports the hypothesis that the motivation does play a key role in the process of critical appraisal. Not just though the effect that motivation has directly on critical appraisal but also indirectly.

Of the four models tested, the re-specified model fits the data the best. This has a number of different implications. First, there are important relationships between all of the different variables that need to be considered when examining critical appraisal. Second, that motivation plays both an important role both directly and indirectly. The re-specified model retains each of these components of the model, and allows for a more parsimonious, better fitting model.

STUDY 2: CROSS VALIDATION OF RE-SPECIFIED MODEL

Post-hoc modification of structural equation models has been criticized (c.f. Cliff, 1983; Cudeck & Browne, 1983; Browne & Cudeck, 1993). The reason for this is that if a model is rejected, then in the view held by some theorists there is nothing more to do other than to reject the model (Byrne, 2001). Post-hoc modification of models, can, and in many cases probably does, take advantage of chance factors, thus making the model functionally useless. In reality, however, most researchers do not reject a model out of hand if the model does not immediately render significant results. Other researchers have claimed that it is all right to modify a rejected model as long as one is aware of the exploratory character of this approach (Byrne, 2001).

In addition to acknowledging the exploratory nature of post-hoc modification of a structural equation model, cross-validation procedures can be used to test the validity of the post-hoc modifications (MacCullum, Roznowski, Necowitz, 1992). For example, in order to determine whether the final model, in this case the re-specified model, is a valid model, the model should be tested on a separate, independent sample. In this way, the re-specified model serves as a calibration sample, and the second, independent, sample serves as a validation sample. The method of cross-validation used is an invariance-testing strategy described by Byrne (2001).

I tested a second, independent sample of undergraduate students using a different domain. The purpose of this procedure was to validate the re-specified model to ensure that the model was not simply capitalizing on chance factors.

Methods

Participants

The validation sample consisted of a total of 205 undergraduate students in introductory psychology classes recruited through a research participant pool. Participants received partial course credit for their participation in this study. Of the 205 participants, 116 were female (approximately 57 %) and 89 were male (approximately 43%). The mean age of participants was 19.0 years, *SD* = 3.00.

Measures and Scoring

The measures and scoring for the validation sample were the same as for the baseline sample except that the domain of interest was eating disorders instead of alcohol abuse. Domain knowledge was assessed with questions 8-42 in Appendix A.Appendix B contains an answer key to the questions that were asked. Cronbach's alpha for the measures were, .58 for domain knowledge, .77 for knowledge of computers and the Web, .61 for knowledge of critical appraisal, and .79 for motivation.

Procedure

Collection of validation sample. The procedure for collecting the validation sample was identical to the procedure used for the calibration sample. Testing occurred in two sessions for each participant.

Cross-validation. To calculate the cross validation, the first step was to create an omnibus, or unconstrained model that contained both the calibration sample and the validation sample. The re-specified baseline model (Figure 25), now forms the basis for this omnibus model. With both of these samples included in the omnibus test, and allowing all of the parameters to vary freely, the omnibus test was then calculated.

Once the omnibus test was calculated, the parameters of the baseline model were then constrained to be equal across both the calibration and the validation samples. This model was then tested and the results of the two models were compared. As described in Byrne (2004), the key measures of interest in this analysis are the chi-square values for the omnibus and the constrained tests. Once these are calculated, the difference in the chi-square values and degrees of freedom for the omnibus and the constrained samples will be used to measure whether there is a significant difference between the two models.

Results

Descriptive Statistics

Differences between baseline and validation samples. There were no significant differences between any of the meausures collected with the baseline sample and those of the validation sample. The only exception to this are self reports of domain knowledge. (see Table 11 for a list of related statistics).

Domain knowledge. Participants were able to answer correctly slightly less than half of the 15 domain related questions (DK1) related to eating disorders (M = 6.52, SD = 1.02). Self-reported domain knowledge (DK2) was 3.03 (out of 5) with a standard deviation of 1.02. Both distributions were approximately normal (see Figures 29 and 30 respectively).

Table 11

Independent-Sample t-tests Between Baseline Measures and Re-specified Baseline Model

Measure	Mean Difference	Standard Error	df	t-statistic	P-value
DK1	-0.250	0.194	413	-1.28	>.05
DK2	0.220	0.094	413	2.35	>.05
CK1	0.256	0.141	413	1.82	>.05
CK2	0.283	0.171	413	1.66	>.05
CK3	0.220	0.153	413	1.44	>.05
KA1	0.084	0.182	413	0.46	>.05
KA2	-0.253	0.130	413	-1.96	>.05
KA3	-0.100	0.145	413	-0.70	>.05
KA4	0.055	0.120	413	0.46	>.05
KA5	-0.264	0.146	413	-1.81	>.05
KA6	-0.106	0.117	413	-0.91	>.05
KA7	-0.046	0.115	413	-0.40	>.05
MO1	1.537	0.788	413	1.95	>.05
MO2	0.323	0.686	413	0.47	>.05
MO3	0.155	0.096	413	1.62	>.05
MO4	0.151	0.099	413	1.52	>.05
CA1	0.076	0.084	413	0.90	>.05

CA2	0.209	0.210	413	1.00	>.05
CA3	-0.008	0.031	413	-0.25	>.05
CA4	-0.178	0.245	413	0.73	>.05
CA5	0.108	0.141	413	0.77	>.05

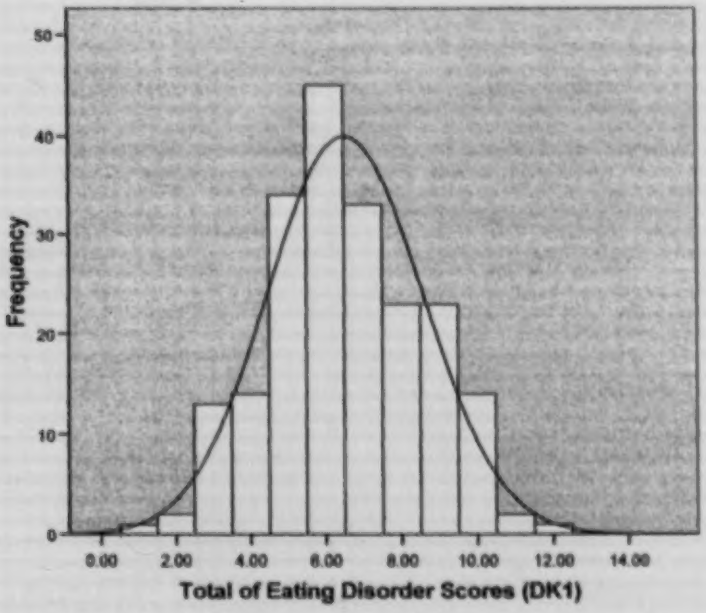

Figure 29. A histogram of the frequency of test of eating disorder domain knowledge (validation sample) with normal curve.

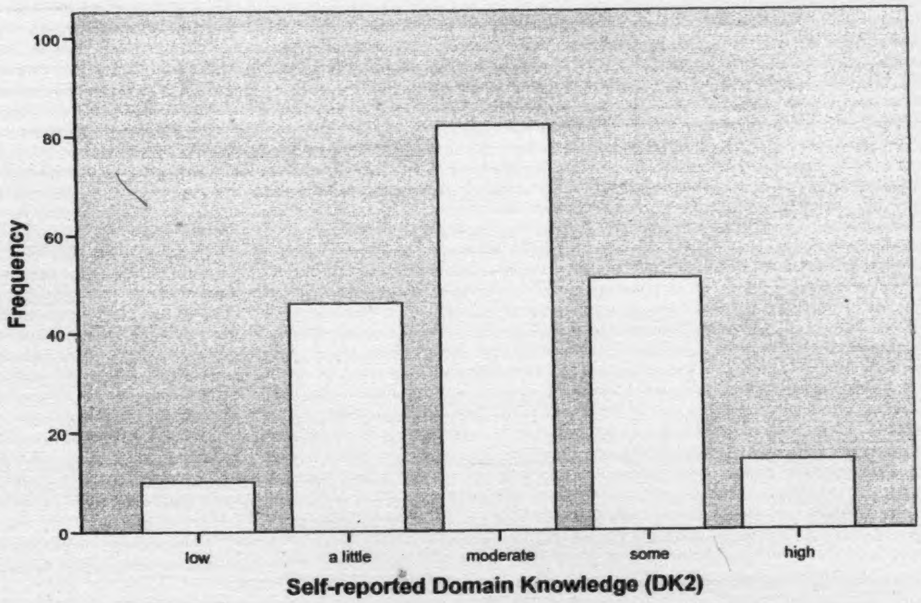

Figure 30. A bar graph of the frequency of self-reported domain knowledge (validation sample).

Knowledge and skill with the Web. The mean response for how much participants liked computers and Web (CK1) was 8.14 (out of total of 10) with a standard deviation of 1.39. The resulting distribution was slightly skewed to the left and was truncated at the top end of the scale see Figure 31).

Questions regarding participants' frequency of use of the computer and the Web (CK2) resulted in a nearly normal distribution (see Figure 32). Mean frequency reported was 10.89 (out of a total of 15) with a standard deviation of 1.79.

The distribution resulting from questions regarding skill with computers/Web (CK3) had a mean of 7.62 and a standard deviation of 1.59 (see Figure 33).

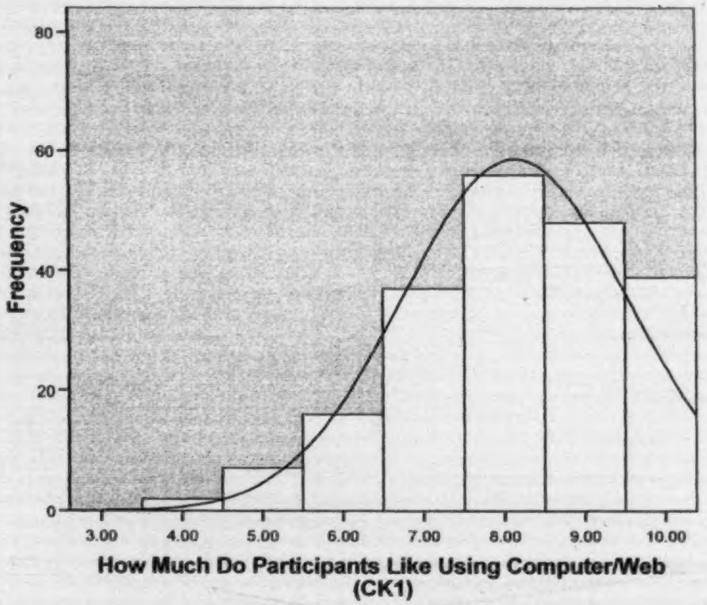

Figure 31. A histogram of the frequency of participants' ratings of how much they like to use Computers and the Web (validation sample) with normal curve.

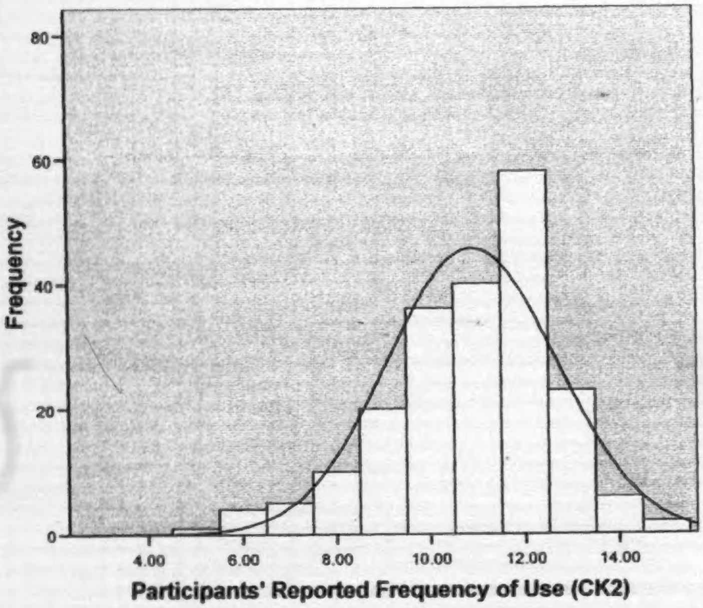

Figure 32. A histogram of the frequency of participants' ratings of how frequently they use Computers and the Web (validation sample) with normal curve.

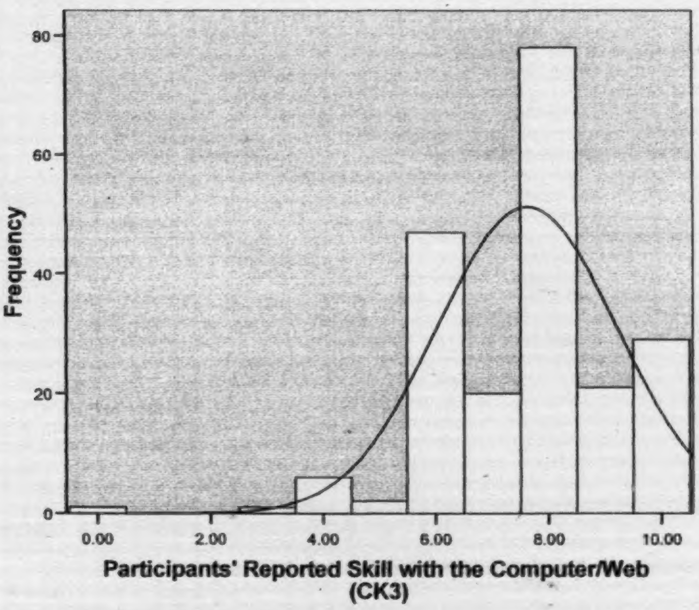

Figure 33. A histogram of the frequency of participants' reported skill with Computers and the Web (validation sample) with normal curve.

Knowledge of critical appraisal. Knowledge of critical appraisal was measured using several different indicators. Each measure had a minimum score of 0 for no valid strategies reported and 4 for each valid strategy reported (higher scores indicating more valid strategies used more frequently). Knowledge of the author (KA1) had a mean of 2.33 and a standard deviation of 1.91 (see Figure 34). Strategies for assessing the credibility/believability of the author (KA2) had a mean of 0.55 and a standard deviation of 1.21 (see Figure 35). Strategies for assessing the purpose of the Web resource (KA3) had a mean of 1.38 and a standard deviation of 1.53 (See Figure 36). Strategies for assessing the accuracy of the Web resource (KA4) had a mean of 0.91 and a standard deviation of 1.29 (see figure 37). Strategies for assessing the objectivity of a Web resource (KA5) had a mean of 1.12 and a standard deviation of 1.33 (see Figure 38). Strategies for assessing the comprehensiveness of a Web resource (KA6) had a mean of 0.63 and a standard deviation of 1.13 (see Figure 39). Strategies for assessing the currency of information found on a Web resource (KA7) had a mean of 1.45 and a standard deviation of 1.18 (see Figure 40).

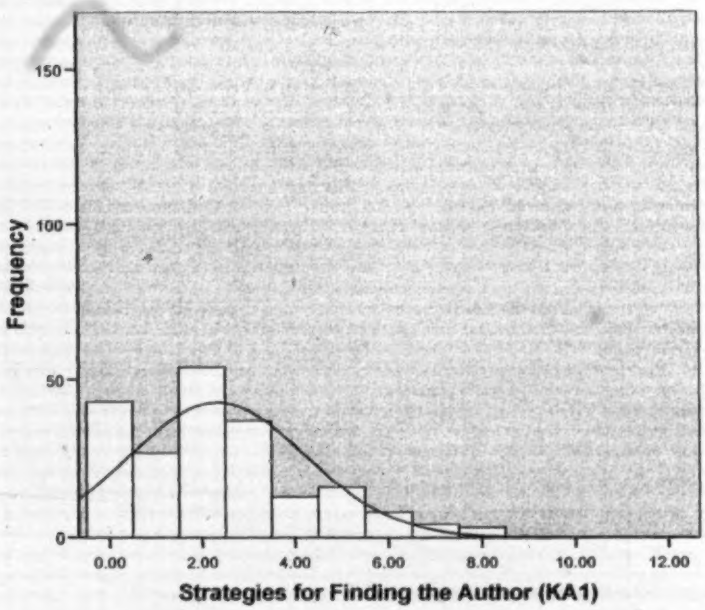

Figure 34. A histogram of participants' strategies for finding the Author of a Web resource (validation sample) with normal curve.

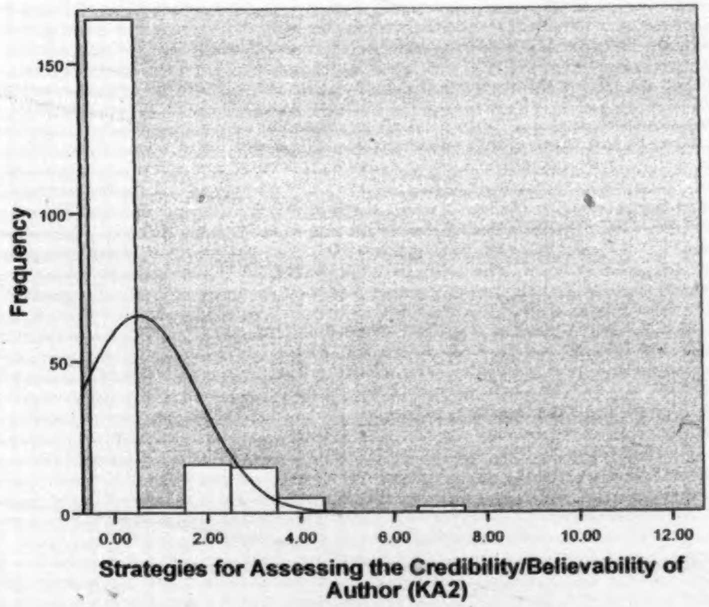

Figure 35. A histogram of participants' strategies for assessing the credibility or believability of the author of a Web resource (validation sample) with normal curve.

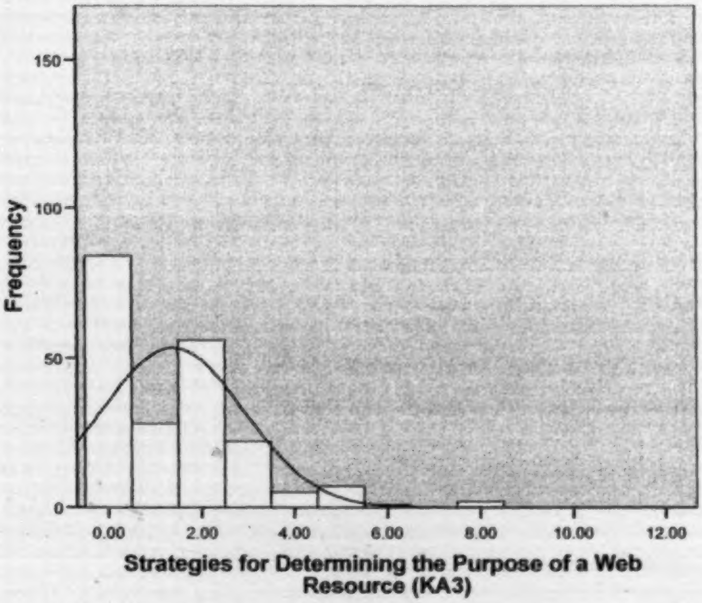

Figure 36. A histogram of participants' strategies for determining the purpose of a Web resource (validation sample) with normal curve.

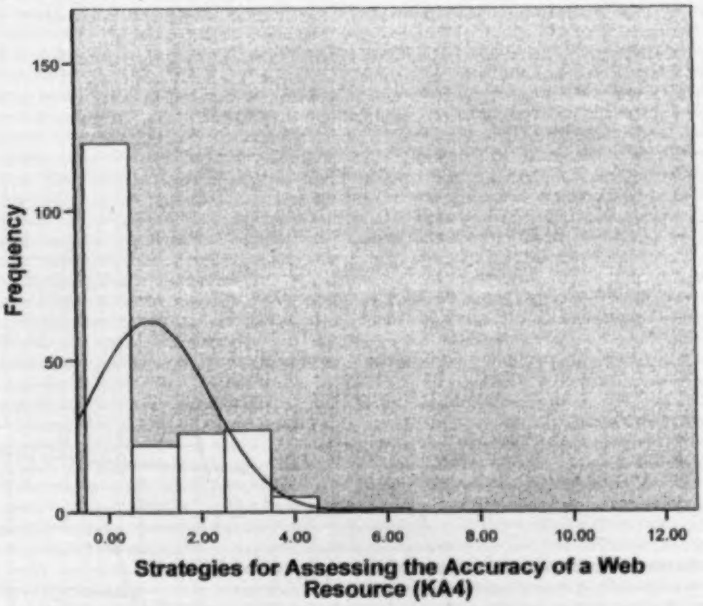

Figure 37. A histogram of participants' strategies for assessing the accuracy of a Web resource (validation sample) with normal curve.

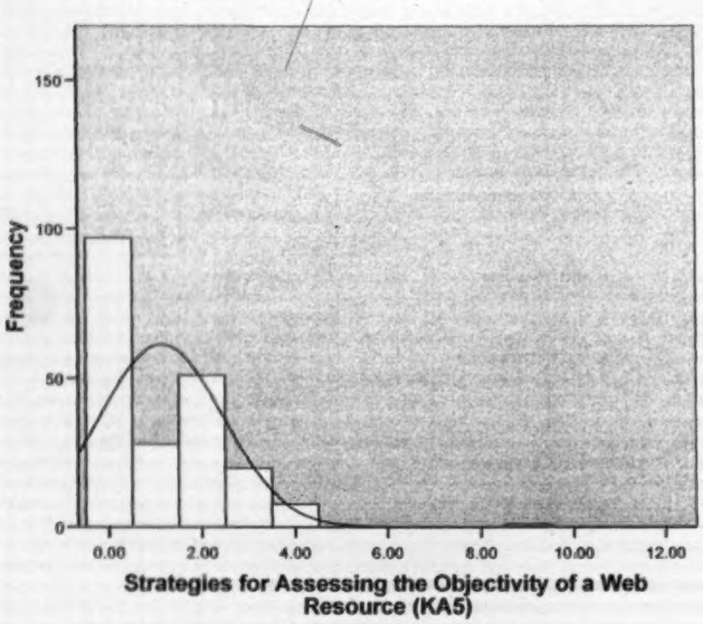

Figure 38. A histogram of participants' strategies for assessing the objectivity of a Web resource (validation sample) with normal curve.

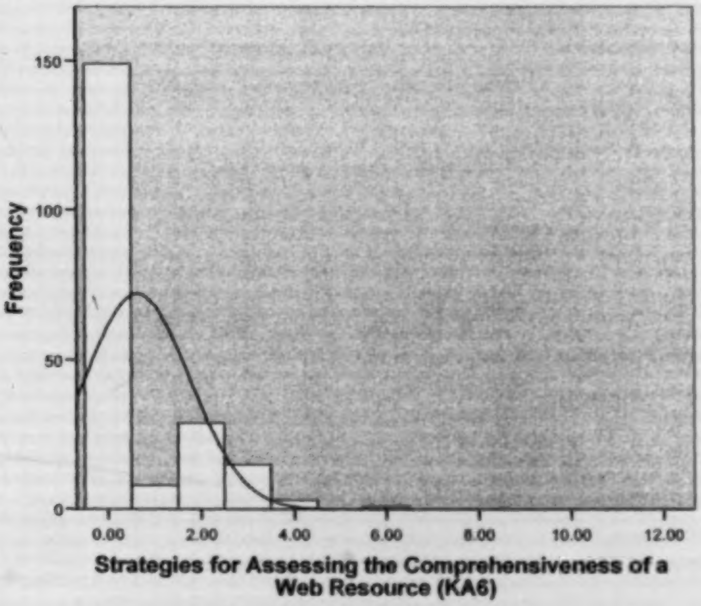

Figure 39. A histogram of participants' strategies for assessing the comprehensiveness of a Web resource (validation sample) with normal curve.

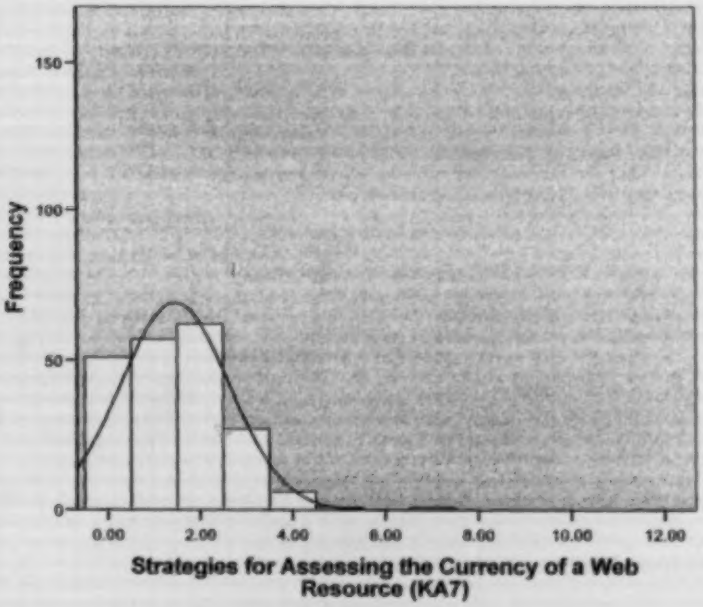

Figure 40. A histogram of participants' strategies for assessing the currency of a Web resource (validation sample) with normal curve.

Motivation. Personal relevance (MO1) scores had a mean of 16.16 (out of a total of 40) with a standard deviation of 9.08 (see Figure 41). Personal responsibility (MO2) had a mean score of 16.98 (out of a possible 35) with a standard deviation of 7.13 (see Figure 42). Motivation today (MO3) had a mean of 3.31 (out of 5) with a standard deviation of 0.93 (see Figure 43). Motivation in general (MO4) had a mean of 2.96 (out of 5) and a standard deviation of 1.02 (see Figure 44).

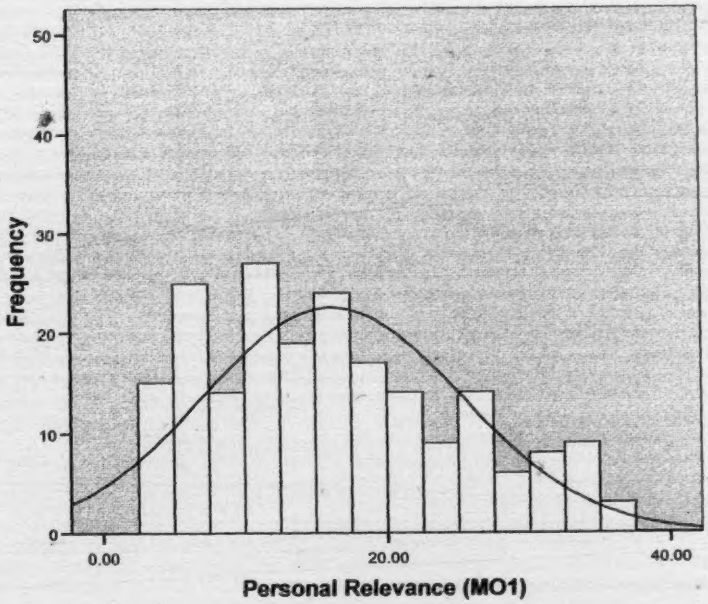

Figure 41. A histogram of participants' ratings of personal relevance (validation sample) with normal curve.

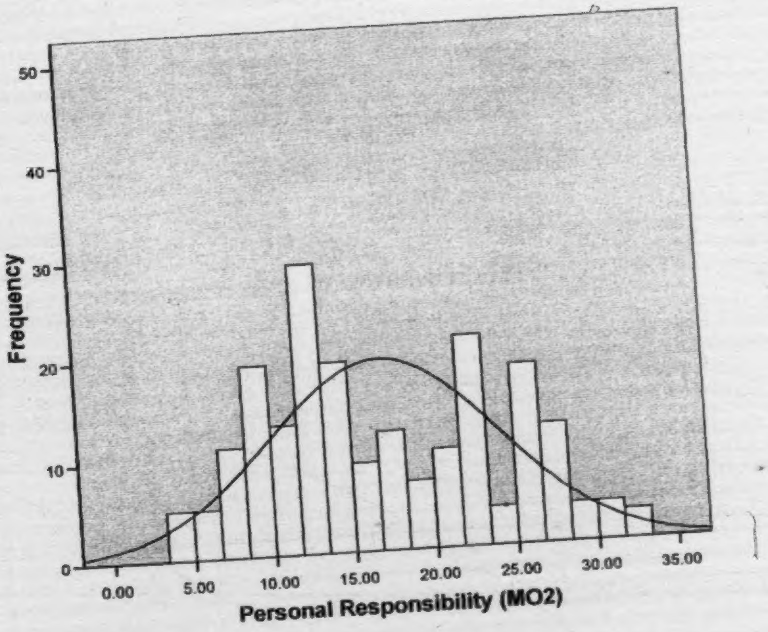

Figure 42. A histogram of participants' ratings of personal responsibility (validation sample) with normal curve.

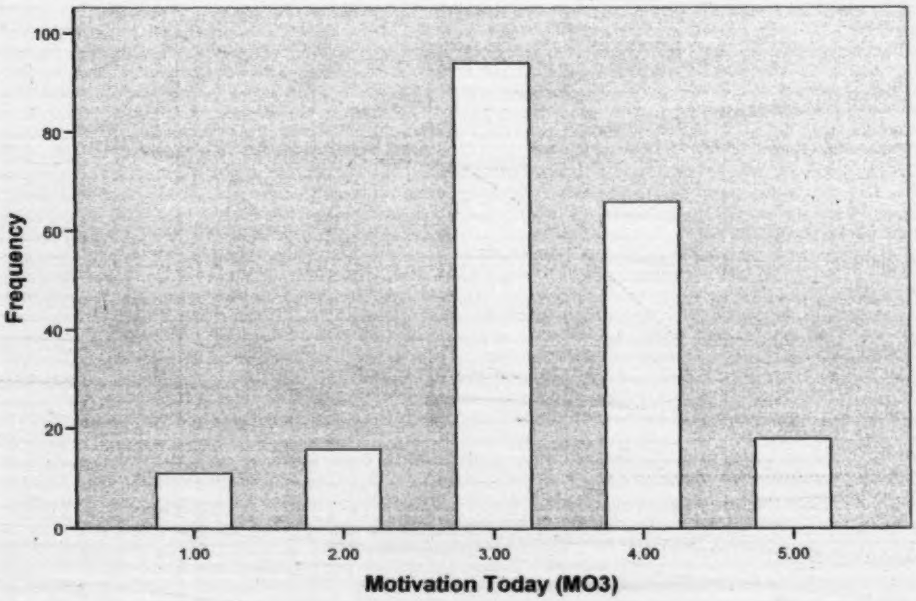

Figure 43. A bar graph of participants' ratings of their motivation on the day of testing (validation sample).

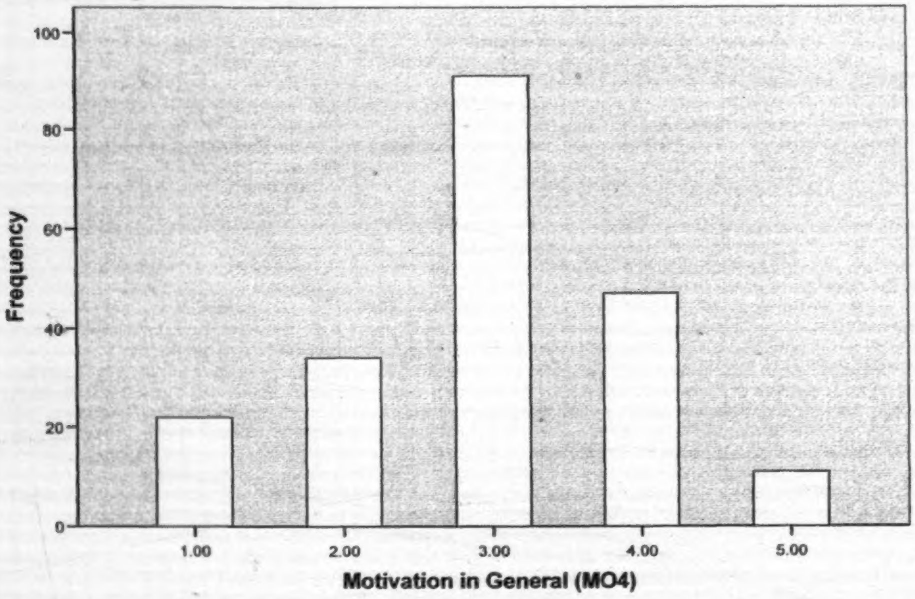

Figure 44. A bar graph of participants' ratings of their general motivation to ensure that information regarding alcohol abuse is accurate (validation sample).

Critical Appraisal. Rating of the Web resource (CA1) was done using a 5-point likert-type scale. Mean response was 3.88 with a standard deviation of 0.80 (see Figure 45). Critical appraisal of the Author (CA2) had a mean of 7.80 (out of 13) with a standard deviation of 2.13 (see Figure 46). Critical appraisal of the purpose of the Web resource (CA3) had only 3 possible values (1, 2, or 3). The mean was 1.07 with a standard deviation of 0.35. Critical appraisal of the content of the resource (CA4) had a maximum score of 12 and the mean response was 7.13 with a standard deviation of 2.51 (see Figure 47). Critical appraisal of the form of the Web resource (CA5) had a mean of 5.56 and a standard deviation of 1.35 (out of a total of 9-see Figure 48).

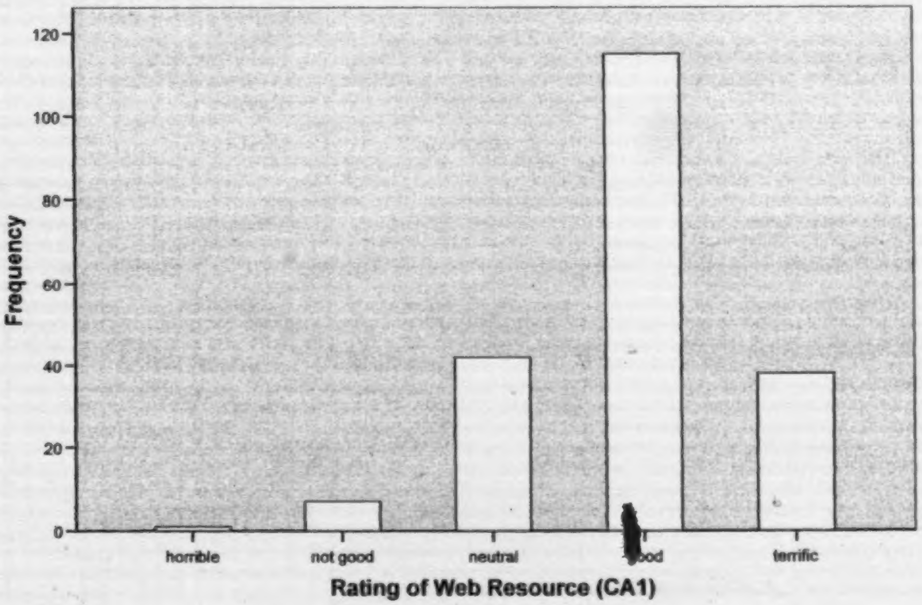

Figure 45. A bar graph of participants' ratings of the test Web resource (validation sample).

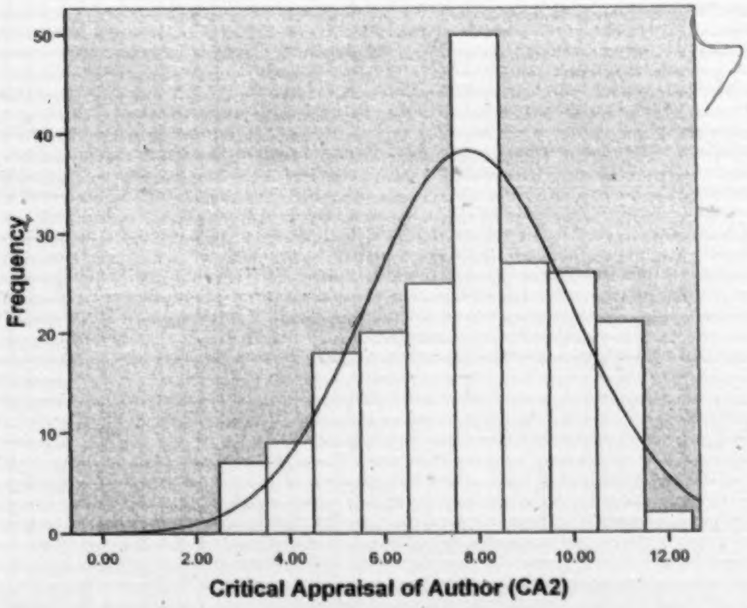

Figure 46. A histogram of participants' critical appraisal scores regarding the author of a Web resource (validation sample) with normal curve.

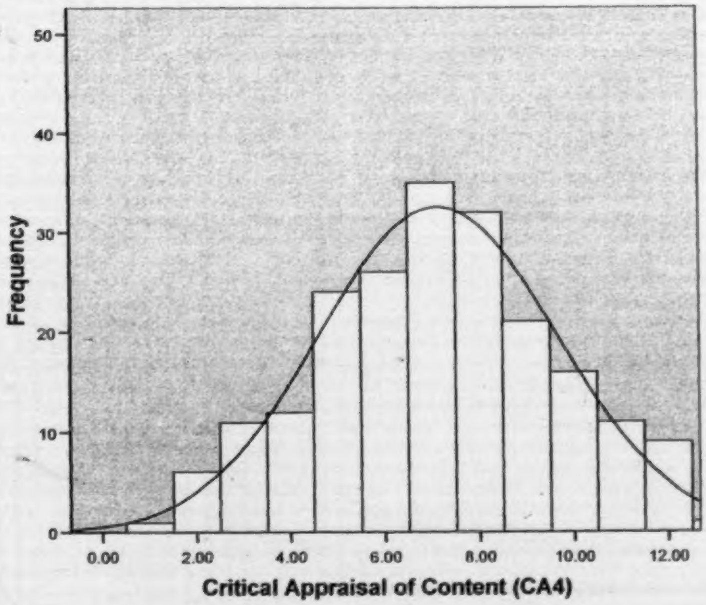

Figure 47. A histogram of participants' critical appraisal scores regarding the content of a Web resource (validation sample) with normal curve.

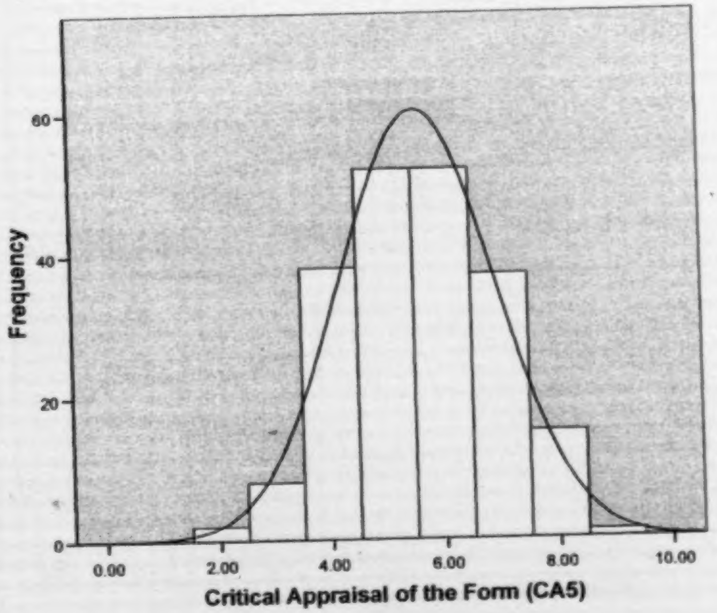

Figure 48. A histogram of participants' critical appraisal scores regarding the form of a Web resource (validation sample) with normal curve.

Omnibus Test

The chi-square value and the associated degrees of freedom for the unconstrained, or omnibus test is ($\chi^2(367, N = 415) = 507.40, p < .05$). See Table 9 for the entire list of fit measures.

Constrained Test

The key values from this analysis are the chi-square value and the associate degrees of freedom ($\chi^2(373, N = 415) = 514.40, p < .05$). See Table 10 for entire list of fit measures.

Cross-validation Measure

The cross-validation measure is calculated by hand by subtracting the unconstrained test chi-square value (507.40) from the constrained test chi-square (514.40). The corresponding degrees of freedom are similarly calculated by subtracting the constrained test degrees of freedom (367) from the unconstrained test degrees of freedom (373). This results in a chi-square difference of 7 with a corresponding degrees of freedom value of 6. This resulting chi-square is not significant ($p > .05$). Comparison of the fit measures shows that the constrained (CFI =0 .72, RMSEA = 0.03) and the omnibus (CFI=0.73, RMSEA =0.03) tests both result in moderately fitting estimates (See Tables 12 and 13).

Table 12

Summary of Fit Indices for the Omnibus Cross Validation

Fit Index	Re-specified Baseline Structural Model
χ^2	507.384
Root Mean Square Residual (RMR)	0.657
Goodness-of-fit Index (GFI)	0.898
Adjusted Goodness-of-fit Index (AGFI)	0.871
Parsimony Adjusted Goodness-of-fit Index (PGFI)	0.713
Normed Fit Index (NFI)	0.574
Relative Fit Index (RFI)	0.512
Incremental Fit Index (IFI)	0.830
Tucker-Lewis Index (TLI)	0.791
Comparative Fit Index (CFI)	0.818
Parsimony Ratio (PRATIO)	0.874
Parsimony-adjusted Normed Fit Index (PNFI)	0.501
Parsimony-adjusted Comparative Fit Index (PCFI)	0.715
Noncentrality Parameter Estimate (NCP)	140.384
NCP lower bound	85.173
NCP upper bound	203.645
Minimum Discrepancy Function (FMIN)	1.229
Population Discrepancy (F0)	0.340
F0 lower bound	0.206
F0 upper bound	0.493
Root Mean Square Error of Approximation (RMSEA)	0.030
RMSEA lower bound	0.024
RMSEA upper bound	0.037

Akaike Information Criterion (AIC)	699.384
Browne-Cudeck criterion (BCC)	722.283
Expected cross validation index (ECVI)	1.693
ECVI lower bound	1.560
ECVI upper bound	1.847
Hoelter .05 index	337
Hoelter .01 index	354

Table 13

Summary of Fit Indices for the Constrained Cross Validation

Fit Index	Re-specified Baseline Structural Model
χ^2	514.439
Root Mean Square Residual (RMR)	0.668
Goodness-of-fit Index (GFI)	0.897
Adjusted Goodness-of-fit Index (AGFI)	0.872
Parsimony Adjusted Goodness-of-fit Index (PGFI)	0.724
Normed Fit Index (NFI)	0.568
Relative Fit Index (RFI)	0.513
Incremental Fit Index (IFI)	0.827
Tucker-Lewis Index (TLI)	0.793
Comparative Fit Index (CFI)	0.816
Parsimony Ratio (PRATIO)	0.888
Parsimony-adjusted Normed Fit Index (PNFI)	0.504
Parsimony-adjusted Comparative Fit Index (PCFI)	0.725
Noncentrality Parameter Estimate (NCP)	141.439
NCP lower bound	85.860
NCP upper bound	205.070
Minimum Discrepancy Function (FMIN)	1.246
Population Discrepancy (F0)	0.342
F0 lower bound	0.208
F0 upper bound	0.497
Root Mean Square Error of Approximation (RMSEA)	0.030
RMSEA lower bound	0.024
RMSEA upper bound	0.036

Akaike Information Criterion (AIC)	692.439
Browne-Cudeck criterion (BCC)	713.668
Expected cross validation index (ECVI)	1.677
ECVI lower bound	1.542
ECVI upper bound	1.831
Hoelter .05 index	338
Hoelter .01 index	354

Discussion

The two samples that were collected were the same with regard to sex, age, education level and all other other factors except for the domain of interest. None of their scores on the common measures that they took are significantly different from each other except for self-reported domain knowledge. This is a good thing, because it means that the only real difference between the groups is the domain which they searched.

Because of the tentative nature of the exploratory phase of model building and testing, there is an increased chance of spurious results. By cross validating the model with a separate sample, I am more confident that the re-specified model is a useful model.

The descriptive statistics of the sample for the validation model are very similar to those of the sample for the baseline model. Again, when examining the domain expertise scores, the scores seem to be clustered about the mid-point. Out of 15 possible, participants were able to answer correctly just under half. This is not surprising considering that this was a non-specialized sample. The resulting normal distribution could describe quite a few different phenomena in the general population. In addition, participants' scores on the test of domain knowledge seem to match very well with their own reported domain knowledge. The scores for the validation model sample are only slightly lower than those of the baseline sample and the values for the self-reported knowledge are almost identical

Results from the analysis of the knowledge of and skill with computers and the Web variables shows these obtained values to be almost identical to the values from the baseline sample. All of the measures (how much they like computers/Web, how frequently the use computers/Web, and how much skill they report having with computers/Web) are on the upper end of the scale.

All of the mean responses for knowledge of critical appraisal variables are still very low. As in the baseline sample, not one participant scored higher than 11, in fact, not one person out of all 205 participants actually scored higher than 9 and the modal response in every one of the indicators was still either 0 or 2.

Motivation scores were in the low to mid range. Personal relevance was slightly higher in the re-specified sample than the baseline sample, but was slightly lower than personal responsibility but both of the responses were near the mid-range. Both motivation today and motivation in general also showed similar patterns.

The rating of the Web resource was positive and critical appraisal of the author was slightly higher than the mid-range. Critical appraisal of the purpose of the Web resource was slightly lower than the mid-range while critical appraisal of the content was slightly higher. Critical appraisal of the form was also fairly high.

The most striking aspect of the descriptive analysis of the validation sample is that the sample is almost identical in most respects to the baseline sample. Thus, one of assumptions being made in the validation of validation sample, that it is they are the same except for the domain, seems to be validated.

Again, as in the baseline sample, domain knowledge seems to be clustered about the mid-point, the knowledge and skill with the Web quite high, the knowledge of critical appraisal for the validation sample seems to be quite low and motivation seems to be slight higher than the mid-range.

The critical appraisal measures do not differ much at all from the values from the baseline sample. All of which supports the contention that the validation sample is an appropriate sample for testing the validity of re-specified model.

As mentioned in the discussion section of the re-specified model, there are some potential concerns when any re-specification of a structural model takes place. Specifically, there is an increased likelihood that any model that does result from re-specification could be the result of chance factors. The cross-validation supports the contention that the re-specified model is a valid model. The fact that there is no significant difference between the chi-square for the unconstrained model and the chi-square for the constrained model is evidence that there are no significant differences in the overall fit of the omnibus and the constrained models.

GENERAL DISCUSSION

The purpose of this study was to examine the relationship between factors that could be important in determining whether critical appraisal will occur when an information seeker assesses information from a health-related Web resource in order to establish a framework. The main findings are that there are a number of interacting factors that need to be considered when examining how critical appraisal of health-related Web resources occurs and that there is sufficient evidence to show that the proposed framework is useful.

Structural equation modeling was chosen as the method of evaluating the connection between the hypothesized predicting variables and the likelihood of critical appraisal. The use of structural equation modeling allowed for an examination of the simultaneous and interacting effects of all of the variables as well as a level of abstraction that is not possible with other analysis techniques. Another benefit of using structural equation modeling is that it allowed for a post-hoc modification of the hypothesized model based on fit. As the primary purpose of this research was model building and testing, this was ideal.

Through the use of structural equation modeling, I determined that the initially hypothesized model was a moderately well fitting model.. As further modifications to the model were made, I was able to create an even better fitting, more parsimonious model. While this new, better fitting model was still significant, the range of fit criteria that were examined led me to the conclusion that the model was, in fact plausible.

I also tested the assumption that in addition to the direct effects of domain knowledge, knowledge of critical appraisal and motivation, there are important covariances between these factors that are also important for critical appraisal. This analysis illustrated that inclusion of covariance measures significantly improved the model fit compared to a model with all measures of covariance removed. As a result, I determined that the re-specifed model was the best fitting model.

A second assumption that I tested in this dissertation was that motivation plays a mediating role with the other factors. I tested a model in which the covariances between motivation and all other factors were removed. I found that this model as significantly less well fitting than then re-specified model so I determined that of the four models that were examined the best fitting was the re-specified model.

Because of the post-hoc modifications that I made to the original model, the chance of spurious results was increased. To ensure that the results that I obtained were not due to chance alone, I used a validation procedure to test the final model with a similar, but independent sample. This second sample was no different from the original baseline sample except that this validation sample was tested using a separate domain. I determined through this validation procedure that there was evidence to suggest that the re-specified, post-hoc modified model that I developed was not due to chance alone. This allows for a certain level of confidence in interpreting the results.

The final, re-specified, and cross-validated model offers insight into the possible processes involved in the determination of why critical appraisal does or does not take place (See Figure 49). The three factors that had a direct impact on

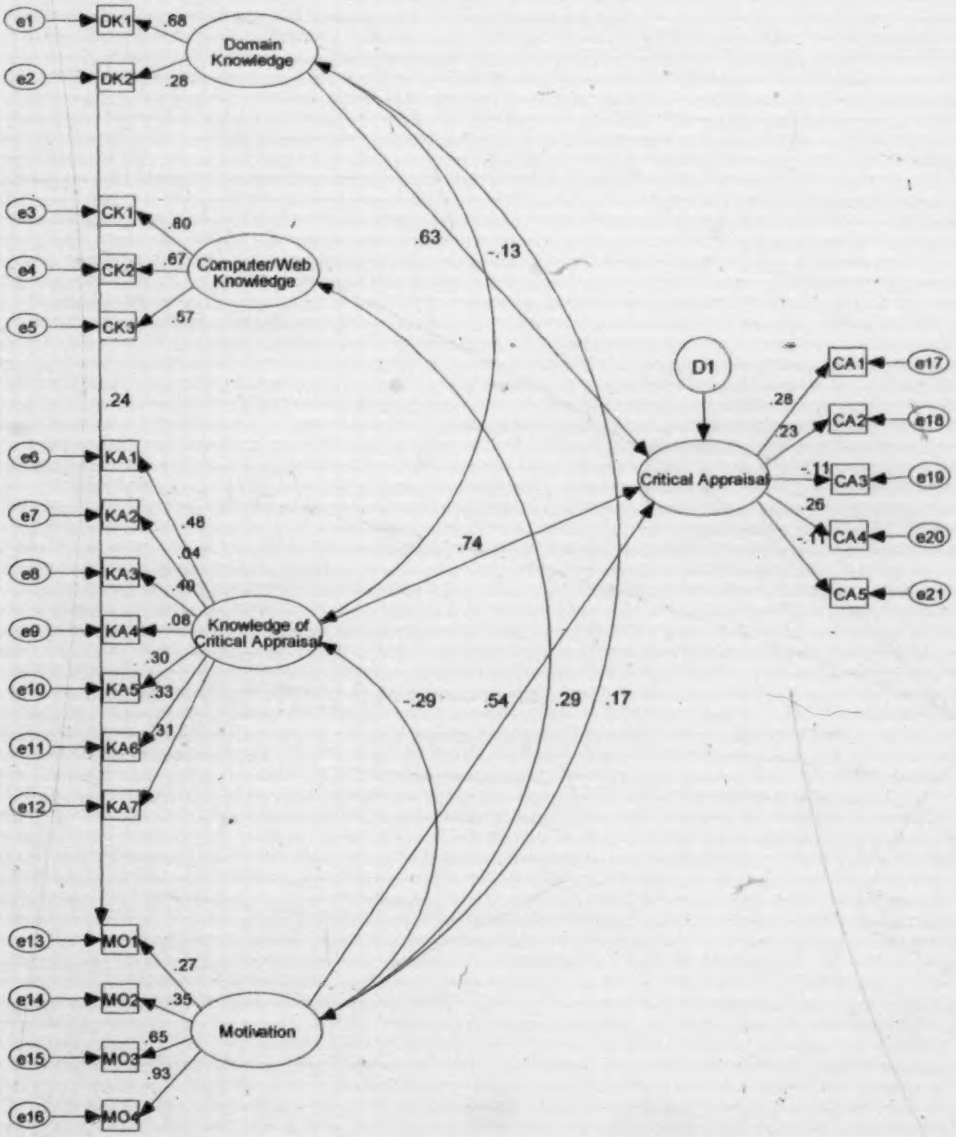
e1
DK1
.68
Domain Knowledge
e2
DK2
.28
e3
CK1
.80
e4
CK2
.67
Computer/Web Knowledge
e5
CK3
.57
.63
-.13
.24
D1
CA1
e17
.28
CA2
e18
.23
Critical Appraisal
-.11
CA3
e19
.26
CA4
e20
-.11
CA5
e21
e6
KA1
e7
KA2
.48
.04
e8
KA3
.40
.74
e9
KA4
.08
Knowledge of Critical Appraisal
.30
e10
KA5
.33
-.29
.54
.29
.17
e11
KA6
.31
e12
KA7
e13
MO1
.27
e14
MO2
.35
.65
Motivation
e15
MO3
.93
e16
MO4

Figure 49. Final re-specified structural model with standardized regression weights.

critical appraisal were domain knowledge, knowledge of critical appraisal, and motivation. In addition to this, although it did not have a direct effect on critical appraisal, computer/Web knowledge does still seem to have an indirect effect through the covariance with motivation. In addition, there were significant measures of covariance between domain knowledge and knowledge of critical appraisal and motivation.

As well, there was a contributing covariance between knowledge of critical appraisal and motivation. Finally there was a direct effect between one of the manifest variables of domain knowledge and one of motivation. Specifically, there was an effect of self-reported domain knowledge on ratings of personal relevance.

As was mentioned, when navigating through the Web, or watching someone navigate through the Web, the speed with which people seem to make decisions regarding the worth of a site and validity of information found on that site are amazing.

When one looks closer, it seems that people are not basing their decisions, in most cases, on any systematic critical analysis, but rather on more non-systematic gut instinct, or visceral approaches. Thus, by examining those things that may be contributing factors, it becomes possible to not only understand why people are not always engaging in critical appraisal, but potentially to find ways to increase the likelihood of critical appraisal occurring.

The Web has provided an increased level of access to information of almost every kind. While this increased access has led to more people finding and using more health related information, there is still no guarantee that the information they find is accurate or that they are able to understand information they come across. The individual information searcher is responsible, in large part, to determine which information and information providers should be believed and then how they should use it.

The samples used in these studies were university students who had recently received instruction in critical appraisal. I expected that these students would score very high in measures of critical appraisal. In contrast, rather, it seems that many of these students avoided engaging in much critical appraisal at all. This is in line with results from Eysenbach and Kohler (2002) who found that although people reported knowing about what they should be looking for in order to critically appraise a Web resource, many still do not do this (c.f., Flanigan & Metzger, 2000).

Different theories have been developed to try to explain the differences between processing on a superficial level and processing on a more deep level (see Wathena & Burkell, 2002 for a review). Most of these tend to rely on the assumption that there is, in fact, a linear relationship between the predictors and the likelihood of engaging in critical appraisal (cf. Petty & Cacioppo, 1986). In addition, there are theories that aim to describe and explain the factors that are important in motivating people to engage in a more effortful process of investigation (c.f., Petty & Cacioppo, 1986).

What most of these theories have in common is the fact that not everyone is equally likely to engage in critical appraisal. There is no way to predict how an individual will critically appraise without knowing how much domain knowledge they have, how much knowledge regarding critical appraisal they have and how motivated they are to engage in critical appraisal.

When all of the potential contributing factors are considered, it leads to a very complex view of critical appraisal, especially regarding decisions about individual resources. As such, some researchers have suggested that studying this level of specificity is not possible in any systematic fashion, nor is it useful (Lazonder, 2000). Rather than leave the question unanswered and treat intra-site decision making as an undiscoverable "black box", this research demonstrates that it is possible to discover and examine some of the subtle factors that may be important for predicting and understanding critical appraisal.

The factors that I studied in this research and that were deemed to be important in explaining why critical appraisal occurs were domain knowledge, knowledge and skill with the Web, and knowledge of critical appraisal and motivation. The motivational factors that I measured were personal relevance, personal responsibility, general motivation, and situational motivation.

The first ability factor related to critical appraisal is domain knowledge. This is, perhaps, not surprising. Adequate domain knowledge may be a contributing factor for a couple of reasons. The first of these is that adequate domain knowledge allows one a context from which one can base a critical appraisal. If a person has little or no domain knowledge, then it is much harder to base a decision regarding the utility of a site, or the veracity of information contained within the site on anything substantial

(Rouet, 2003).

The second reason why domain knowledge may play such an important role is that it may actually reflect an underlying interest in the subject matter. This may also be seen through the connection between self-reported domain knowledge and personal relevance. If people are interested in a particular domain, they are more likely to be motivated to ensure that the information that they are using is accurate (Lazonder, Biemans, & Wopereis, 2000). Thus, domain has both a direct influence on the likelihood of critical appraisal and a moderating effect through its influence on other, potential, contributors.

The second ability factor that contributes directly to critical appraisal is knowledge of critical appraisal. The more that someone knows about critical appraisal, the more likely it is that they will engage in critical appraisal. What this indicates is that critical appraisal may not be something that is spontaneously engaged in. Research by Taylor et al. (2000) has also shown that there is a connection between training on critical appraisal and the use of critical appraisal strategies.

On a merely intuitive level, this seems to make some sense as well. In order for one to engage effectively in a process, it is usually necessary to have some knowledge, or skill with the process in question. For example, the likelihood of one driving a car is directly related to the person's knowledge of how to drive a car. If people were asked to self-report on whether they can drive or not, would anyone be surprised if those who reported that they can drive are more likely than those who report that they cannot drive to actually drive a car? While it true that driving skill alone is not a sufficient explanation for predicting the likelihood of driving, because even if you know how to drive you may decide not to, it is a necessary explanation. In other words, it is conspicuous by its absence. If you know how you can either choose to or choose not to. If you do not know how, then there is no choice. In a similar manner, simply knowing about critical appraisal may not make one more likely to critically appraise, but in order to critically appraise; one does have to have a certain, minimum level of knowledge.

Knowledge of critical appraisal is related to both domain knowledge and motivation.It is not readily apparent why domain knowledge and motivation would be related, but a possible explanation may have to do with the fact that there could be some other third variable that is influencing both motivation and knowledge of critical appraisal. What is, perhaps, most interesting about this connection is that it is a negative relationship. This is difficult to interpret unless we posit a separate variable influencing both of these factors. Such third, or extraneous variables could be things like personality characteristics, socio-economic status, situational variables or any number of factors that could influence a person's level of knowledge about what should be done, and the motivation to actually engage in the behaviors.

In addition to ability factors, motivation to critically appraise is also important. One of the most difficult aspects of this research was to find domains which interested participants enough to convince them to put some effort into the research task. One may have the necessary skill to critically appraise, but unless one also possesses the proper motivation, this skill will probably not be demonstrated. For example, in experiments on latent learning in rats, researchers were able to show that rats did learn in mazes without any reinforcement, but that the rats would not demonstrate this learning unless motivated to do so (Whishaw, 1999). So too with humans, skills can be mastered but not demonstrated unless there is sufficient reason to do so. Motivation moderates the relationship between ability and critical appraisal.

With the multiple, competing demands for time and cognitive resources during everyday tasks, one must be selective in the allocation of these resources. Research by Chin and Brown (2000) on levels of processing shows that while deep processing (an example of which critical appraisal could be considered) does lead to a better understanding of what is being processed, it also requires a significant investment in terms of cognitive resources (see also: Craik & Tulving 1975). That being said, it is not possible to devote cognitive resources to every task in equal amounts (Navon & Gopher, 1979). Thus what often has to happen is that the processor has to decide where to devote cognitive resources.

When searching the Web, there may be many times when it is important to ensure that information is correct and that the Web resource being used is reputable. However, there may also be times when it does not matter. If accuracy really does not matter for the person who is viewing a site, it not as likely that critical appraisal will take place. In health-related domains, however, it is often vitally important to ensure that information is accurate, especially if health-related decisions are being based upon that information.

Some of the motivating factors that were hypothesized to play a role in determining whether someone would decide whether to expend the resources necessary for critical appraisal are personal relevance, personal responsibility, general motivation and situational motivation. These factors are all important parts of explaining whether motivation will be a contributing factor in determining whether critical appraisal will take place. Lack of motivation also helps to explain why many people, including many of the participants in this research do not engage in critical appraisal to the extent that they potentially could.

The separate parts of the model that play a role in predicting or at least explaining critical appraisal all seem to work together to explain critical appraisal. This model is an important first step in trying to discover what the underlying components of critical appraisal are.

This model allows for a starting point for many different lines of research. This is the first step in a larger program of critical appraisal research. Now that this model has been validated it may be possible to look at different applications. For example, it may be possible to look at how children critically appraise and how this skill develops over time. It may also be possible to look at differences between professionals and non-professionals. Each of these questions, and many more can be studied within the context of this model in a much more controlled and rigorous manner.

Now that I have that the role that motivation plays a role in predicting critical appraisal and moderating the influence of other ability-related factors, it may also be possible to vary motivation levels and to examine the corresponding effect on decision making. What this can do is to allow for a better understanding of the cognitive processes involved in using the Web.

Additionally, it is possible to expand the examination of critical appraisal of the Web beyond health-related domains to determine how domain influences the likelihood of critical appraisal. As Metzger (…) has shown, domain is an important indicator of the amount of critical appraisal that may take place. This model was developed specifically for health-related domains but may also inform critical appraisal in other domains.

The decision making process extends beyond the decision about whether or not the information is useful or not to the decision involving whether one will even engage in a critical appraisal process. So, for those who are interested in the use of critical appraisal, this is important research because it does lead to conclusions about what are the most significant factors in understanding how people are critically appraising health-related information that they find(Chi, 2002).

While the concern over the quality of information on the Web is valid, what to do about it is not so straightforward. The answer cannot be to simply do away with the Web or to try to increase the barriers to access. There is ample evidence from the past to show that this approach will not actually be beneficial. While it is true that there is a lot of inaccurate information on the Web, heath related and non-health related, there is also much good and useful information on the Web. The same issues of inaccuracy exist in non-Web information sources like newspapers and textbooks. The same thing also applies to information that is found on the television and the radio. In addition to being able to critically appraise, people also need to feel a need to critically appraise in order to ensure that information is accurate. This is something that needs to begin long before university. For example, children should be taught critical appraisal skills in elementary school. They should be taught not to blindly accept information from any source, not textbooks, not teachers, not the Media, and not the Web (Korpan, Bisanz, Bisanz, & Henderson, 1997).

The temptation to use each instance of failure to critically appraise as evidence of an inability on the part of the information seeker should be resisted. As shown in this research, there is more involved than ability. There may be times when a person has the ability to critically appraise but chooses not to expend the mental effort necessary for critical appraisal. Critical appraisal is a necessary skill that needs to be taught, but simply possessing the skill is no guarantee that critical appraisal will or should take place. Recognizing Web users as adaptive, efficient processors of information and understanding why people do or do not critically appraise is essential.

In order to critically appraise, you have to have the ability to do so. This ability can come in many different forms. The ones that were discussed here are domain knowledge and knowledge of critical appraisal.

As was demonstrated in this study, trying to get people to critically appraise is not an easy task. Even though the participants of this study had a fairly high level of theoretical knowledge related to critical appraisal (i.e., they knew about what they should be looking for), they did not all engage in critical appraisal. One important caveat that needs to be made regarding the results of this study, and the sample that was used is that this was a sample of university undergraduates. In theory they were a select sample who should have more knowledge of critical appraisal, especially because they were all introductory psychology students and had recently been instructed about the process of critical appraisal. So if critical appraisal was not universal with the samples used in this study, chances are that it is even less likely in the general population.

While it is not practical to expect that everyone can or will become a domain expert in every possible health-related field, some knowledge may be easier to help information seekers to learn about critical appraisal. It is not enough to expect people to just pick up these skills as they make their way through the educational system. Children need to be taught at very young ages about critical appraisal, about what information they need to make decisions. Even the undergraduates who were part of this study, while probably more knowledgeable than the general public regarding the process of critical appraisal, still did not all engage in critical appraisal.

People need to be taught to think for themselves and become discriminating consumers of information so they will not have to rely as much on filters, censorship, or on authority figures. If people are not taught to critically appraise information, from any source, their lack of ability will continue to be a bounding condition and will continue to decrease the likelihood that they will engage in critical appraisal.

In addition to helping information seekers to develop necessary skills, they also need to be taught to understand the importance of critical appraisal. Even the best skills will be of little use if information seekers are not motivated to engage in critical appraisal. If they do not see the importance or the utility of critical appraisal, it is unlikely that they will devote energy and cognitive resources to such a task. How do people develop a desire to critically appraise? Again, the answer has to be education. If people are made aware or taught the importance of critical appraisal it may increase their motivation to do so.

In this study, I found that people were somewhat knowledgeable about critical appraisal, yet they still did not actually critically appraise. Thus, the importance of motivation is manifest. Just as with knowledge of critical appraisal and domain knowledge. It is important that people know about the importance of critical appraisal.

In the end, it is really up to information seekers to decide how much energy they will devote to the process of critical appraisal. No one method is going to immediately cause people who are not engaging in critical appraisal to suddenly do so, however, the likelihood of this happening more often will increase. Ability factors make it possible for critical appraisal to occur, but it is motivation that dictates whether this will happen.

In addition, not every situation will necessitate critical appraisal. Remembering that humans are efficient processors (Simon, 1957), it is not surprising that sometimes even with appropriate skill it is possible to actively choose not to critically appraise.

The Web is continuing to increase in its reach and in the amount of information available. While the format of the Web may change, it is unlikely that access to information will actually decrease. The issues, dilemmas, and problems that are important with the Web right now will only continue to increase in importance as the Web continues to spread in influence. While many of these issues share their roots with other phenomena, the Web has truly revolutionized the manner in which information can be accessed.

One of the primary goals of this study was to establish a framework to situate research on the critical appraisal of heath-related information on the Web. While there are some limitations of the study, the overall objective was achieved. The resulting model can be used as a basis for research in many different fields and will hopefully help to add to our understanding of how and why people use technology such as the Web in the manner that they do.

REFERENCES

Alexander, P. A., Kulikowich, J. M., & Schulze, S. K. (1994). How subject-matter knowledge affects recall and interest. *American Educational Research Journal*, *31*, 313–337.

Arbuckle, J. L. (1994). AMOS: Analysis of moment structures. *Psychometrika, 59,* 135-137.

Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999). Conceptualizing critical thinking. *Journal of Curriculum Studies, 31,* 285-302.

Barfield, W. (1986). Expert-novice differences for software: Implications for problem solving and knowledge acquisition. *Behaviour and Information Technology, 5,* 15-29.

Bentler, P.M. & Bonet D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588-606.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238-246.

Boechler, P. M. (2001). How spatial is hyperspace? Interacting with hypertext documents: Cognitive processes and concepts. *CyberPsychology & Behavior, 4,* 23-46.

Browne M. W. & Cudeck R. (1993). Alternative ways of assessing model fit in K.A. Bollen & J.S. Long (Eds.), *Testing Structural Equation Models*, (pp. 445-455). Newbury Park, CA: Sage.

Browne, M. N., Freeman, K. E., & Williamson, C. L. (2000). The importance of critical thinking for student use of the Internet. *College Student Journal, 34,* 391-399.

Bush, V. (1945). As we may think. *Atlantic Monthly, 176,* 101 -108.

Bucklin, R. E. & Sismeiro, C. (2003). A model of Web site browsing behavior estimated on clickstream data. *Journal of Marketing Research, 40,* 249-267.

Byrne, B. (2001). *Structural Equation Modeling with AMOS: Basic Concepts, Applications and Programming.* Mahwah, NH: Lawrence Erlbaum Associates.Canadian Media Research Corporation. (2005).

Canada Online: A Comparative Analysis of Intenet Users and Non-users in Canada and World: Behaviour, Attitudes and Trends 2004. Toronto: Canadian Internet Project.

Chase, W. G. & Simon, H. A. (1973). Perceptions in chess. *Cognitive Psychology, 4,* 55-81.

Chen, F., Bollen, K. A., Paxton, P., Curran, P. J., & Kirby, J. B. (2001). Improper solutions in structural equation models: Causes, consequences, and strategies. *Sociological Methods and Research, 29,* 468-508.

Chen, S. & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-Process Theories in Social Psychology* (pp. 73-96). New York, NY: Guilford Press.

Chi, E. H. (2002). Scent of the Web. In: Julie Ratner (ed.), *Human Factors and Web Development*, pp. 265-285. Hillsdale, New Jersey: Erlbaum.

Chin, C. & Brown, D. E. (2000). Learning in science: A comparison of deep and surface approaches. *Journal of Research in Science and Teaching, 37,* 109-138.

Claypool, H. M., Mackie, D.M., Garcia-Marques, T., McIntosh, A. & Udall, A. (2004). *Social Cognition, 22(3),* 310-335.

Cole, C. & Leid, J.E. (2003). Using the user's mental model to guide the integration of information space into information need. *Journal of the American Society for Information and Technology, 54,* 39-46.

Conklin, J. (1987). Hypertext: An introduction and survey. *Computer, IEEE Computer Society, 20(9), 17-41.*

Craik F. I. M., & Tulving E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104,* 268-294.

Cudeck, R., & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research, 18,* 147-167.

D'Alessandro, D. M., Kingsley, P., & Johnson-West, J. (2001). The readability of pediatric patient education material on the World Wide Web. *Archives of Pediatric and Adolescent Medicine, 155,* 807-812.

Diamantopoulos, A. & Siguaw, J. A. (2000). *Introducing LISREL: A Guide for the Uninitiated.* London: Sage.

Eisenstein, E. (1979). *The printing press as an agent of change.* Cambridge: Cambridge University Press.

Eveland, W. P., & Dunwoody, S. (1998). Users and navigation patterns of a science World Wide Web site for the public. *Public Understanding of Science, 7*, 285-311.

Eveland, W. P., & Dunwoody, S. (2000). Applying research on the uses and cognitive effects of hypermedia to the study of the World Wide Web. *Communication Yearbook, 25,* 79-113.

Eysenbach, G. & Köhler, C. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus group, usability tests, and in-depth interviews. *British Medical Journal, 324,* 573-537.

Eysenbach, G., Powell, J., Kuss, O., & Sa, E. (2002). Empirical studies assessing the quality of health information for consumers on the World Wide Web: A systematic review. *Journal of the American Medical Association, 287,* 2691-2700.

Fazio, R. H. & Towles-Schwen, T. (1999). The MODE model of attitude-behavioral processes. In S. Chaiken & Y. Trope (Eds.), *Dual-Process Theories in Social Psychology* (pp. 97-116). New York, NY: Guilford Press.

Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of Internet information credibility. *Journalism & Mass Communication Quarterly, 77*(3), 515-540.

Flanagin, A. J., & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of Web-based information. *New Media & Society, 9*(2), 319-342.

Fogg, B. J., Soohoo, C., Danielson, D., Marable, L., Stanford, J., & Tauber, E. R. (2002). How do people evaluate a Web site's credibility? Results from a large study. *Consumer Web Watch*. Retrieved October 21, 2004, from http://www.consumerwebwatch.org/news/report3_credibilityresearch/stanfordPTL.pdf

Fox, S. (2006). Online health search. *Pew Internet & American Life Project*. Washington DC. Retrieved March 5, 2007, from http://www.pewinternet.org/pdfs/PIP_Online_Health_2006.pdf

Graber, M.A., Roller, C.M. & Kaebler, B. (1999). Readability levels of patient education materials on the World Wide Web. *Journal of Family Practice, 48,* 58-61.

Goth, G. (2005). Who-and where-are the new media gatekeepers. *IEEE Distributed Systems Online, 6.* Retrieved April 12, 2007 from http://csdl2.computer.org/comp/mags/ds/2005/07/o7003.pdf

Hair, J. F., Tatham, R. L., Anderson, R. E., & Black, W. (1998). *Multivariate Data Analysis*. London: Prentice-Hall International.

Hargittai, E. (2002). Beyond logs and surveys: In-depth measures of people's Web use skills. *Journal of the American Society for Information Science and Technology, 53,* 1239-1244.

Hansen, D., Derry, H., Resnick, P., & Richardson, C. (2003). Adolescents searching for health information on the Internet: An observational study. *Journal of Medical Internet Research, 5(4),* e25. retrieved from June 20, 2004 from http://www.jmir.org/2003/4/e25/

Hassebrock, F., Johnson, P. E., Bullemer, P., Fox, P. W., Moller, J. H. (1993). When less is more: Representation and selective memory in expert problem solving. *American Journal of Psychology, 106,* 155-189.

Havelock, E. (1982). *The Literate Revolution in Greece and its Cultural Consequences*. Princeton, NJ: Princeton University Press.

Hayduk, L., Mha, X. & Carter-Snell, C. (2002). *Structural Equation Modeling and hierarchical linear modeling: Communicating across disciplines.* Edmonton, AB: Research Laboratory of Department of Sociology, University of Alberta.

Hertzum, M., Andersen, H. H. K., Andersen, V. & Hansen, C.B. (2002). Trust in information sources: Seeking information from people, documents, and virtual agents. *Interacting with Computers*, 14(5), 575-599.

Hill, A. & Spittlehouse, C. (2003). What is critical appraisal? *Hayward Medical Communications.* Retrieved September 25, 2004, from http://www.evidence-based-medicine.co.uk

Hodkinson, C., Kiels, G. & McColl-Kennedy, J. (2000). Consumer web search behaviour: diagrammatic illustration of wayfinding on the web. *International Journal of Human-Computer Studies, 52,* 805-830.

Horigan, J. B. (2006). The Internet as a resource for news and information about science: The convenience of getting scientific material on the web opens doors to better attitudes and understanding of science. *Pew Internet & American Life Project.* Washington DC. Retrieved March 3, 2007, from http://www.pewinternet.org/pdfs/PIP_Exploratorium_Science.pdf

Hoyle, R. H. (1995). The Structural Equation Modeling Approach: Basic Concepts and Fundamental Issues. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues and Applications* (pp. 1-15). Thousand Oaks, California. Sage Publications Inc.

Hu, L. & Bentler, P. M. (1995). Evaluating Model Fit. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues and Applications* (pp. 37-62). Thousand Oaks, California. Sage Publications Inc.

Hu, L. -T & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6,* 1-55.

Huberman, B. A., Pirolli, P. L. T., Pitkow, J. E., & Lukose, R. (1998). Strong regularities in World Wide Web surfing. *Science, 280,* 95-97.

Jenkins, C., Corritore, C.L. & Wiedenbeck, S. (2003). Patterns of information seeking on the Web: A qualitative study of domain expertise and Web expertise. *Information Technology & Society, 1,* 64-89.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling: Second Edition.* New York: Guilford Press.

Kline, R. B. (2006). Reverse arrow dynamics: formative measurement and feedback loops. In G. R. Hancock & R. O. Mueller (Eds.) *Structural Equation Modeling: A Second Course* (pp. 43 – 68). Greenwood CT: Information Age Publishing Inc.

Korpan, C. A., Bisanz, G. L., Bisanz, J., & Henderson, J. M. (1997). Assessing literacy in science: Evaluation of scientific news briefs. *Science Education, 81,* 515-532.

Lazonder, A.W. (2000). Exploring novice users training needs in searching information on the WWW. *Journal of Computer Assisted Learning, 16,* 326-325)

Lazonder, A.W., Biemans, H. J. A. and Wopereis, I.G.-J.H. (2000). Differences between novices and experienced users in searching information on the World Wide Web, *Journal of the American Society for Information Science, 51,* 576-581.

MacCullum, R. C. (1995). Model Specification: Procedures, Strategies and Related Issues. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues and Applications* (pp. 16-36). Thousand Oaks, California. Sage Publications Inc.

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology, 66,* 37-47.

Martin, H. (1994). *The history and power of writing.* (L. G. Cochrane, Trans.) Chicago, The University of Chicago Press. (Original work published 1988)

Mayer, M. & Till, J. E. (1996). The Internet: A modern Pandora's box? *Quality of Life Research, 5,* 568-571.

Metzger, M. J., Flanagin, A. J., & Zwarun, L. (2003). College student Web use, perceptions of information credibility, and verification behavior. *Computers & Education, 41,* 271-290.

Metzger, M. J. (2005). Understanding how Internet users make sense of credibility: A review of the state of our knowledge and recommendations for theory, policy, and practice. In R. Weingarten, & M. Eisenberg (Eds.), *Proceedings of the Internet Credibility and the User Symposium* [online]. Seattle, WA. http://www.ischool.washington.edu/credibility/Metzger%20Skills.pdf

Metzger, M. J. (2006). Effects of site, vendor, and consumer characteristics on Web site trust and disclosure. *Communication Research, 33*(3), 155-179.

National Library of Medicine. (1998). National Library of Medicine to Work with Public Libraries to Help Consumers Find Answers to Medical Questions. National Library of Medicine Press Release. 1998 Oct 22. Retrieved February 23, 2006, from http://www.nlm.nih.gov/archive/20040831/news/press_releases/medplus.html

Navon, D. & Gopher, D. (1979). On the economy of the human processing system. *Psychological Review, 86,* 214-255.

Niederhauser, D. S., Reynolds, R. E., Salmen, D. J., & Skolmoski, P. (2000). The influence of cognitive load on learning from hypertext. *Journal of Educational Computing Research, 23,* 237-255.

Olsen, D. R. (1994). *The world on paper: The conceptual and cognitive implications of writing and reading.* Cambridge: Cambridge University Press.

Rideout, V. (2001). *How young people use the Internet for health information.* Menlo Park: Henry J. Kaiser Family Foundation.

Petty, R. E. & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed), *Advances in Experimental Social Psychology Volume 19.* (pp. 124-205).Orlando, Fl: Academic Press, Inc.

Petty, R. E. & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual-Process Theories in Social Psychology* (pp. 41-72). New York, NY: Guilford Press.

Randel, J. M. & Pugh, H. L. (1996). Differences in expert and novice situation awareness in naturalistic decision making. *International Journal of Human-Computer Studies,* 579-597.

Rouet, J. (2002). What was I looking for? The influence of task specificity and prior knowledge on students' search strategies in hypertext. *Interacting with Computers, 15,* 409-428.

Rothman, A. J., & Schwartz, N. (1998). Constructing perceptions of vulnerability: Personal relevance and the use of experiential information in health judgments. *Personality and Social Psychology Bulletin, 24*, 1053-1064.

Sanbonmatsu, D. M. & Fazio, R. H. (1990). The role of attitudes in memory-decision making. *Journal of Personality and Social Psychology, 59,* 614-622.

Sawasdichai, N. & Poggenpohl, S. (2003). User Analysis: Framework. *Visible Design, 37.*

Schuette, R. A. & Fazio, R. H. (1995). Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model. *Personality and Social Psychology Bulletin, 21,* 704-710.

Shuyler, K. S. & Knight, K. M. (2003). What are patients seeking when they turn to the Internet? Qualitative content analysis of questions asked by visitors to an orthopaedics Web site. *Journal of Medical Internet Research 2003; 5(4),* e24. Retrieved June 20, 2004 from http://www.jmir.org/2003/4/e24/

Simon, H. A. (1957). *Models of Man.* New York, NY: John Wiley & Sons.

Stanford, J., Tauber, E., Fogg, B. & Marable, L. (2002). Experts vs. Online consumers: A comparative credibility study of health and finance Web sites. *Consumer Web Watch.* Retrieved October 21, 2004, from http://www.consumerwebwatch.org/news/report3_credibilityresearch/slicedbread.pdf

Tang, H., Ng, J. H. K. (2006). Googling for a diagnosis-use of Google as a diagnostic aid: Internet based study. British Medical Journal, doi:10.1136/bmj.39003.640567.AE

Taylor, R., Reeves, B., Ewings, P., Binns, S. Keast, J. Mears, R. (2000). A systematic review of the effectiveness of critical appraisal skills training for clinicians. *Medical Education, 34,*120-125.

Varnhagen, C. K. (2002). *Making sense of Psychology on the Web.* New York, NY: Worth Publishers.

Venezky, R. L. (1991). The Development of Literacy in the Industrialized Nations of the West. In R. Barr, M. L. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of Reading research, Volume II* (pp. 46-67). White Plains, NY: Longman Publishing Group.

Vicente, K. J. (1992). Memory recall in a process control system: A measure of expertise and display effectiveness. *Memory and Cognition, 20,* 356-373.

Vicente, K. J. & Wang, J. H. (1998). An ecological theory of expertise effects in memory recall. *Psychological Review, 105,* 33-57.

Wathen, C. N. & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology, 52,* 134-144.

Whishaw, I. Q. (1991). Latent learning in a swimming pool place task by rats: evidence for the use of associative and not cognitive mapping processes. *The Journal of Experimental Psychology. B, Comparative and Physiological Psychology, 43,* 83-103.

Zhu, B., Walter, S. D., Rosenbaum, P. L., Russell, D. J., & Raina, P. (2006). Structural equations and log-linear modeling: a comparison of methods in the analysis of a study on caregivers' health. *BMC Medical Research Methodology, 6:49.* Retrieved May 17,2007, from http://www.biomedcentral.com/1471-2288/6/49.

APPENDIX A

1. How would you classify your level of knowledge with respect to alcohol abuse and alcoholism?

1 Low	2	3	4	5 High

2. How would you classify your level of knowledge with respect to eating disorders?

1 Low	2	3	4	5 High

3. For the average adult a safe level of drinking is:

   a) up to two drinks* per day for men and one drink* per day for women and older people
   b) up to two drinks* per day for men and women and one drink* per day for older people
   c) up to three drinks* per day for men and two drinks* per day for women and older people
   d) up to three drinks* per day for men, two drinks* per day for women and none for older people

4. True or False: Alcohol is a leading factor in all three of the top causes of death of people between the ages of 15 and 25.

   a) True
   b) False

5. How does alcohol abuse differ from alcoholism?

   a) There is no difference, these are the same.
   b) The effects of alcoholism are, in general, more severe than alcohol abuse.
   c) Alcohol abuse does not involve intense craving or physical dependence
   d) Alcoholism can be treated, but alcohol abuse cannot.

6. How long does it take for alcohol to affect the brain?

   a) 10 seconds
   b) 90 seconds
   c) 10 minutes

d) 90 minutes

7. How does a family history of alcoholism affect someone's risk of being an alcoholic?

a) Children of alcoholics cannot drink at all or they will become alcoholics.
b) Genes can cause vulnerability to alcoholism.
c) There is no added risk of alcoholism to children of alcoholics.
d) Growing up in a home with alcoholics can cause vulnerability to alcoholism even in adopted children
e) Both genes and growing up in a home with alcoholics affect risk.

8. How can you tell if you are sober enough to drive safely?

a) If you can recite the alphabet backwards, touch your finger to your nose with your eyes closed, and walk a straight line.
b) If you feel sober, you probably are.
c) If an hour has passed since your last drink*.
d) If an hour has passed for each drink* you had.

9. True or False: Alcohol affects women differently than men.

a) True
b) False

10. With treatment, people diagnosed as alcoholics can expect to:

a) be cured and go on to lead a relatively normal, independent life
b) remain relatively unchanged, alcoholism cannot be treated
c) make vast improvements, but never truly be cured
d) none of the above

11. How long does it take the body to process or break down one drink*?

a) Half an hour
b) One hour
c) An hour and a half
d) Depends on the drink*

12. Which of the following may result from prolonged drinking?

a) liver disease
b) heart disease
c) certain forms of cancer
d) pancreatitis
e) all of the above

13. What types of treatment are most effective at curing alcoholism?

a) Psycho-pharmacological treatments.
b) Psychotherapy.
c) Intensive Behavioral therapy.
d) None of the above

14. A safe level of drinking for someone who is pregnant is:

a) up to the equivalent of two drinks* per week.
b) no more than the equivalent of one drink* per week
c) up to the equivalent of two drinks* per if alcohol is accompanied by a large meal.
d) there is no safe level of drinking if you are pregnant.

15. According to research done in the United States, what percentage of college students meets the criteria for a diagnosis of alcohol abuse?

a) 11%
b) 67%
c) 31%
d) 48%

16. Which of the following is a drug that is now used to help people remain sober by reducing the craving for alcohol?

a) Naltrexone
b) Librium
c) Disulfiram
d) Valium

17. Which of the following is not a symptom of alcohol poisoning?

- a) Person may be unconscious but cannot be awakened.
- b) Uncontrollable shaking
- c) Breathing is slow or irregular.
- d) Vomiting while passed out and not waking up during or after vomiting.
- e) Cold, clammy, unusually pale or bluish skin

18. __% of females and __% of males suffer from an eating disorder

- a) 5, 1
- b) 1, 5
- c) 3, 2
- d) 4, 2

19. The onset of __% of eating disorders occurs during adolescence

- a) 95
- b) 50
- c) 75
- d) 85

20. Which of the following is not a major characteristic of all eating disorders?

- a) Disturbed body image
- b) Intense fear of weight gain and becoming fat
- c) Relentless obsession to become thinner
- d) Significantly low body weight

21. Anorexia nervosa is characterized by a body weight of __% of normal body weight.

- a) 70
- b) 85
- c) 60
- d) 75

22. Eating disorders most often strike:

- a) young women between 12 and 25, across all socio-economic and ethnic lines
- b) overweight people
- c) athletes
- d) young professionals, who are on the "fast track"

23. Which of the following is not usually found in anorexic patients?

a) Low heart rate and blood pressure
b) Osteoporosis
c) Brain abnormalities
d) Vitamin and mineral deficiencies

24. Bulimia is treated by:

a) Restrictive diet and exercise
b) Antipsychotic medication
c) Antidepressant medication
d) Laxatives

25. Bulimics are usually

a) Severely underweight
b) Underweight to normal weight
c) Normal weight to overweight
d) Severely overweight

26. Which of the following statements about men and eating disorders is false?

a) A substantial proportion are homosexual, bisexual or asexual
b) They are more likely to have been obese before the onset of the disorder
c) There is a stronger association between dieting and sports activity in men than in women
d) They make up 15-20% of all eating disorder sufferers

27. True or false: Most eating disorders are caused by a desire to lose weight that has gone to extremes.

a) True
b) False

28. Who most likely suffers from an eating disorder?

- a) Cynthia, a 42-year-old housewife, waits until her husband and kids leave for work or school and then raids the fridge and pantry, eating ice cream, donuts, potato salad, leftover lasagna, and half a bag of cookies before stopping.

- b) Larry, a young man of 21, tries not to eat so he can feel some power over food. He often makes gourmet meals for his friends but doesn't partake along with them, even though he's recently lost more than 20 pounds and friends comment about how thin he is. He also enjoys finagling dinner invitations so he can turn them down.

- c) Jan, a college student, goes out for pizza and eats five pieces. She feels bad about it, so she sneaks to the bathroom and throws up. This is not the first time she's done this it seems to help her keep from gaining weight despite her poor eating habits.

- d) all of the above

29. True or False: Anorexia nervosa and bulimia are mental illnesses.

- a) True
- b) False

30. Which of the following personality characteristics is not usually associated with anorexics?

- a) Highly emotionally reserved and cognitively inhibited
- b) Excessive focus on perfectionism
- c) Show reduced conformity and respect for others
- d) Avoid risks and prefer routine and predictable environments

31. Which of the following is not a familial characteristic leading to anorexic behavior?

- a) Parental over direction of the child
- b) A lack of emphasis on propriety and rule-mindedness
- c) Limited tolerance of disharmonious affect or psychological tension
- d) Poor skills in conflict resolution

32. Which of the following is not true about bulimics?

a) They are preoccupied with shame and guilt about their behavior
b) They are in denial about the seriousness of their disorder
c) They struggle painfully to master the urge to binge
d) They go to great lengths to conceal their bulimic behavior

a

APPENDIX B

Answer Key for Test of Domain Knowledge

Alcoholism/Alcohol Abuse	Eating Disorders
Question 3 -A	Question 18-A
Question 4 -A	Question 19-D
Question 5 -C	Question 20-D
Question 6 -B	Question 21-B
Question 7 -E	Question 22-A
Question 8 -D	Question 23-D
Question 9 -A	Question 24-C
Question 10-C	Question 25-C
Question 11-B	Question 26-D
Question 12-E	Question 27-B
Question 13-D	Question 28-D
Question 14-C	Question 29-A
Question 15-C	Question 30-C
Question 16-A	Question 31-B
Question 17-B	Question 32-B

APPENDIX C

1. Do you have a computer with Internet connection at home?

1 2

Yes No

2. How much do you like using the computer?

1 2 3 4 5

Not at all Very much

3. How much do you like searching the Web?

1 2 3 4 5

Not at all Very much

4. How frequently do you use the computer?

1 2 3 4 5

Not at all Very frequently

5. How frequently do you search the Web?

1 2 3 4 5

Not at all Very frequently

6. How frequently do you search for health-related information?

1 2 3 4 5

Not at all Very frequently

7. How skilled are you using the Web?

1 2 3 4 5

Not at all Very skilled

8. How skilled are you searching the Web?

1 2 3 4 5

Not at all Very skilled

APPENDIX D

For each of the following questions, please indicate all the different methods/strategies (not just the ones that you might use) that could be used to evaluate some aspect of a Web resource, and then for each one indicate how frequently you use the methods/strategy.

43. What are different ways to determine the author of an informational Web resource?

Way to determine the author	How frequently do you do this?			
	All the time	Frequently	Infrequently	Never
1.	1	2	3	4
2.	1	2	3	4
3.	1	2	3	4

44. How do you determine **how credible or believable the author** of an informational Web resource is?

Way to determine credibility/believability of the author	How frequently do you do this?			
	All the time	Frequently	Infrequently	Never
1.	1	2	3	4
2.	1	2	3	4
3.	1	2	3	4

45. What are different ways to determine the **purpose** of an informational Web resource? Write as many ways as you can think of to determine the purpose and then indicate how often you use each method.

Way to determine the purpose	How frequently do you do this?			
	All the time	Frequently	Infrequently	Never
1.	1	2	3	4
2.	1	2	3	4
3.	1	2	3	4

46. What are different ways to determine the **accuracy** of an informational Web resource?

Way to determine the accuracy	How frequently do you do this?			
	All the time	Frequently	Infrequently	Never
1.	1	2	3	4
2.	1	2	3	4
3.	1	2	3	4

47. How do you determine how **objective** an information Web resource is?

Way to determine objectivity	How frequently do you do this?			
	All the time	Frequently	Infrequently	Never
1.	1	2	3	4
2.	1	2	3	4
3.	1	2	3	4

48. How do you determine how **comprehensive** an information Web resource is?

Way to determine comprehensiveness	How frequently do you do this?			
	All the time	Frequently	Infrequently	Never
1.	1	2	3	4
2.	1	2	3	4
3.	1	2	3	4

49. How do you determine how **current** an information Web resource is?

Way to determine currency	How frequently do you do this?			
	All the time	Frequently	Infrequently	Never
1.	1	2	3	4
2.	1	2	3	4
3.	1	2	3	4

45. What are different ways to determine the **purpose** of an informational Web resource? Write as many ways as you can think of to determine the purpose and then indicate how often you use each method.

Way to determine the purpose	How frequently do you do this?			
	All the time	Frequently	Infrequently	Never
1.	1	2	3	4
2.	1	2	3	4
3.	1	2	3	4

46. What are different ways to determine the **accuracy** of an informational Web resource?

Way to determine the accuracy	How frequently do you do this?			
	All the time	Frequently	Infrequently	Never
1.	1	2	3	4
2.	1	2	3	4
3.	1	2	3	4

47. How do you determine how **objective** an information Web resource is?

Way to determine objectivity	How frequently do you do this?			
	All the time	Frequently	Infrequently	Never
1.	1	2	3	4
2.	1	2	3	4
3.	1	2	3	4

48. How do you determine how **comprehensive** an information Web resource is?

Way to determine comprehensiveness	How frequently do you do this?			
	All the time	Frequently	Infrequently	Never
1.	1	2	3	4
2.	1	2	3	4
3.	1	2	3	4

49. How do you determine how **current** an information Web resource is?

Way to determine currency	How frequently do you do this?			
	All the time	Frequently	Infrequently	Never
1.	1	2	3	4
2.	1	2	3	4
3.	1	2	3	4

APPENDIX E

Criteria for Assessing Validity of Critical Appraisal Strategies

43.

Way to determine the author
1.Check top/bottom for author's name
2.Check for links that might help you identify the author
3.Try stripping of parts of the URL to get a directory or subdirectory

44.

Way to determine credibility/believability of the author
1. Conduct Web search for author using common search engine (e.g., Google, Yahoo!, etc.)
2. Conduct search using specialized internet directory (e.g., Switchboard, Yahoo! People, etc.)
3. Conduct bibliographic search on author using PsychINFO or ERIC (i.e., look for publications)

45.

Way to determine the purpose
1. Use search engine to determine who is linking to the site
2. Look for statement of purpose
3. Look for information (clues with respect to the purpose) on the site

46.

Way to determine the accuracy
1. Reputable source
2. Look for evidence of peer review
3. Check copyright statements or links that could represent sponsorship (i.e., look for bias)

47.

Way to determine objectivity
1. Check copyright statements or links that could represent sponsorship
2. Look for a statement of purpose
3. Determine whether the site is trying to sell something
4. Determine whether the site is trying to convince people of something
5. Look for corroborating evidence on other sites
7. Check for references

48.

Way to determine comprehensiveness
1. Conduct search for similar resources
2. Collect multiple resources
3. Follow up on links or bibliographic info in the resource

49.

Way to determine currency
1. Look for date (posted, last updated, copyrighted) on page
2. Check Page Infor/Properties and look for last modified date
3. Strip off the resource location to get a directory or subdirectory that might tell you when the resource was last modified
4. Search for the title on AltaVista
5. Assess the number of outdated links

APPENDIX F

Please indicate whether the following statements are applicable to you.

27. Do you know someone with an alcohol abuse problem?

1 2

Yes No

28. Do you have a family member or friend who has an alcohol abuse problem?

1 2

Yes No

29. Do you feel that alcohol abuse has a direct effect on your life?

1 2 3 4 5

Not at all Very much

30. Do you feel that alcohol abuse is personally relevant to you?

1 2 3 4 5

Not at all Very much

31. Is alcohol abuse a topic that is personally interesting to you?

1 2 3 4 5

Not at all Very much

32. Do you ever search the Internet for information about alcohol abuse?

1 2 3 4 5

Not at all Very much

33. If you search the Internet for information about alcohol abuse, do you search:

a) for your own interest's sake?

1 2

Yes No

b) to make a health-related decision for yourself?

1 2

Yes No

c) to make a health-related decision for someone else?

1 2

Yes No

d) to communicate with someone else?

1 2

Yes No

e) to recommend a site for someone else?

1 2

Yes No

34. Relative to other information you search for, do you feel it is *as* important to ensure that information you read about alcohol abuse is accurate?

1 2 3 4 5

Not at all **Very much**

35. Generally, when you search for information on alcohol abuse, do you feel a personal sense of responsibility to ensure that information you read is accurate and reliable?

1 2 3 4 5

Not at all **Very much**

36. If you were to evaluate a Web resource related to alcohol abuse, would you be concerned that people would judge you based on your evaluation of the resource?

1 2 3 4 5

Not at all **Very much**

37. Are you concerned that you will be judged today?

1 2 3 4 5

Not at all **Very much**

38. **Today**, an accurate assessment of your motivation to critically appraise Web resources related to alcohol abuse is:

1 2 3 4 5

Not at all **Very much**

39. **In general**, an accurate assessment of your motivation to critically appraise Web resources related to alcohol abuse is:

1	2	3	4	5
Not at all				Very much

APPENDIX G

Evaluation of Web Resource

1. On a scale from 1-5, how would you rate this Web resource?

1	2	3	4	5
Horrible, I would not recommend this site to anyone				Terrific, I would definitely recommend this site to someone else

Please indicate which of the following criteria your evaluation of the Web resource was based on:

AUTHOR (The person(s) who created the content of the Web resource)

2. The identity of the author(s) of the Web resource is clear.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

3. The credentials of the author(s) of the Web resource are listed.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

4. The affiliations of the author(s) of the Web resource are listed.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

5. The author(s) of the Web resource are qualified to create content for the topic.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

6. The identity of the host of the Web resource is clear.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

7. All advertisements are clearly distinguishable from the content of the Web resource.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

8. The host of the Web resource discloses possible conflicts of interest.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

9. A reputable agency, institution, or organization hosts the Web resource.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

10. The Web resource is free from corporate sponsorship *(e.g., pharmaceuticals company)*.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

11. The Web resource is free from commercial advertising.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

12. The links within the Web resource are clearly labeled.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

13. The links within the Web resource function properly.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

14. Registration or membership is necessary to use the Web resource.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

15. Anyone may use the Web resource.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

PURPOSE (To inform, educate, persuade, advocate, entertain, sell, or some combination of these)

16. The purpose of the Web resource is clear.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

17. The Web resource is appropriate for the targeted user group.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

CONTENT (The accuracy, objectivity, comprehensiveness, and currency of the information contained in the Web resource)

18. References for content are provided in the Web resource.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

19. The content of the Web resource does not contain any unsubstantiated claims about treatments.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

20. The content of the Web resource is based on clinical or scientific studies.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

21. The content of the Web resource is presented in an unbiased manner.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

22. The content of the Web resource has been reviewed by experts.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

23. General disclaimers for the use of content are noted in the Web resource.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

24. The content of the Web resource contains a wide range of references.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

25. Information about the range of treatment options is provided in the Web resource.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

26. In-depth information about specific treatments is provided in the Web resource.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

27. Links to other relevant Web resources are provided in the Web resource.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

28. The creation date of the Web resource is noted.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

29. The last update of the Web resource is noted.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

30. The content on the Web resource is current.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

FORM (The physical features of the Web resource)

31. The Web resource is organized in a logical manner.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

32. The Web resource contains an internal search engine.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

33. The Web resource contains mechanisms for feedback.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

34. The Web resource contains contact information for questions or concerns.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

35. The Web resource is easy to navigate.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

36. The multimedia (i.e., graphics, audio, and video) of the Web resource serve a clear purpose.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

37. The Web resource can be personalized to provide only the information you request.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

38. The Web resource can be used by special needs groups (i.e. visually and hearing impaired).

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

39. The Web resource is visually appealing.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

40. The content of the Web resource is free of typographical errors.

1	2	3
No, I didn't think to look for this information	No, this is not important information to know	Yes, this is important information to know

APPENDIX H

Please indicate whether the following statements are applicable to you.

27. Do you know someone with an eating disorder?

1 2

Yes No

28. Do you have a family member or friend who has an eating disorder?

1 2

Yes No

29. Do you feel that eating disorders have a direct effect on your life?

1 2 3 4 5

Not at all Very much

30. Do you feel that eating disorders are personally relevant to you?

1 2 3 4 5

Not at all Very much

31. Are eating disorders a topic that is personally interesting to you?

1 2 3 4 5

Not at all Very much

32. Do you ever search the Internet for information about eating disorders?

1 2 3 4 5

Not at all Very much

33. If you search the Internet for information about eating disorders, do you search:

a) for your own interest's sake?

1 2

Yes No

b) to make a health-related decision for yourself?

1 2

Yes No

c) to make a health-related decision for someone else?

1 2

Yes No

d) to communicate with someone else?

1 2

Yes No

e) to recommend a site for someone else?

1 2

Yes No

34. Relative to other information you search for, do you feel it is *as* important to ensure that information you read about eating disorders is accurate?

1 2 3 4 5

Not at all Very much

35. Generally, when you search for information on eating disorders, do you feel a personal sense of responsibility to ensure that information you read is accurate and reliable?

1 2 3 4 5

Not at all Very much

36. If you were to evaluate a Web resource related to eating disorders, would you be concerned that people would judge you based on your evaluation of the resource?

1 2 3 4 5

Not at all Very much

37. Are you concerned that you will be judged today?

1 2 3 4 5

Not at all Very much

38. **Today**, an accurate assessment of your motivation to critically appraise Web resources related to eating disorders is:

1 2 3 4 5

Not at all Very much

39. **In general**, an accurate assessment of your motivation to critically appraise Web resources related to eating disorders is:

1	2	3	4	5
Not at all				Very much